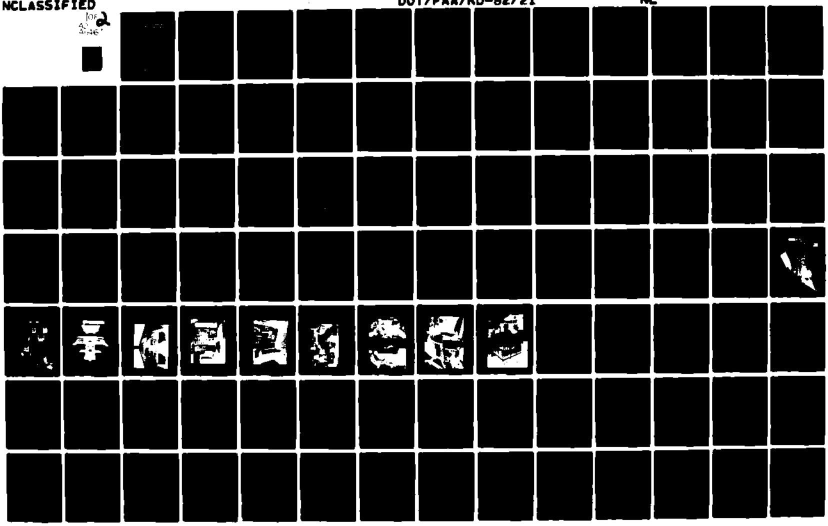
AD-A114 167 WRIGHT STATE UNIV DAYTON OH SCHOOL OF MEDICINE F/G 1/3
FLIGHT CREWMEMBER WORKLOAD EVALUATION.(U)
APR 81 R L SULZER, W J COX, S R MOHLER F33615-80-K-3627
UNCLASSIFIED DOT/FAA/RD-82/21 NL

1 OF 2
AD A114167



(14)

DOT/FAA/RD-82/21
ASF-82-1

Systems Research &
Development Service
Washington, D.C. 20590

Flight Crewmember Workload Evaluation

R. L. Sulzer
W. J. Cox
S. R. Mohler

Wright State University
School of Medicine
P.O. Box 927
Dayton, Ohio 45401

AD A114167

April 1981

Final Report

This document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161.

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/RD-82/21 DOT/FAA/ASF-82/1	2. Government Accession No. AD-A114 167	3. Recipient's Catalog No.
4. Title and Subtitle FLIGHT CREWMEMBER WORKLOAD EVALUATION		5. Report Date April 1981
		6. Performing Organization Code
		8. Performing Organization Report No.
7. Author's) Richard L. Sulzer, William J. Cox and Stanley R. Mohler		
9. Performing Organization Name and Address Wright State University School of Medicine P.O. Box 927 Dayton, Ohio 45401		10. Work Unit No. (TRAIS)
		11. Contract or Grant No. F 33 615 80 K 3627
		13. Type of Report and Period Covered Final August 1980 - April 1981
12. Sponsoring Agency Name and Address U. S. Department of Transportation Federal Aviation Administration Systems Research and Development Service Washington, D.C. 20591		
		14. Sponsoring Agency Code ARD-300
15. Supplementary Notes		

16. Abstract

This is a report on transport category airplane flight crew workload measurement techniques as used in cockpit development and aircraft certification tests by major U.S airframe manufacturers. It reviews the fundamentals of: crew size certification; workload measures and criteria; workload studies made during aircraft design; and workload studies made after the design has been established, including those used in flight test. Certain documentation practices are identified. The limitations of the currently used practices and the needs for improved workload measurement techniques are addressed.

17. Key Words		18. Distribution Statement
Workload	Crew Procedures	Document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161.
Crew Size	Cockpit Functions	
Crew Tasks	Cockpit Mockups	
Simulators	Flight Deck	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	107	

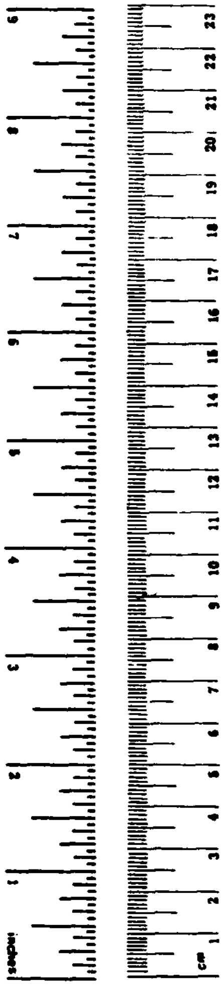
Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

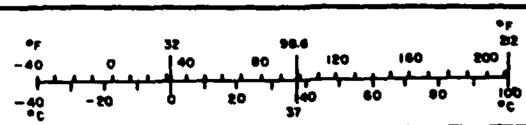
Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	*2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	square meters	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	milliliters	ml
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	liters	l
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C

*1 in ≡ 2.54 (exactly). For other exact conversions and more detailed tables, see NBS Misc. Publ. 286, Units of Weights and Measures, Price $2.25, SD Catalog No. C13.10:286.



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	centimeters	0.4	inches	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	square kilometers	0.4	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F



PREFACE

This report has been prepared to describe, clarify, and catalog specific flight crewmember workload determination techniques which have been used in the past. The information may be useful to type certification applicants in planning and conducting evaluations during stages of aircraft development that will lead to adequate documentation that crew workload is appropriate.

Nothing herein should be construed as establishing a rigid formula for current or future transport aircraft cockpit certification programs. Such programs are individually developed, appropriate tests being selected for critical or changed features of the particular aircraft, and a modified test program designed each time. Over the past two decades, however, there have been a number of such individually designed certification programs from which much can be learned. These programs have been successful as evidenced by actual line service in both domestic and world wide operations. This report reviews the laboratory, engineering, simulation, and flight test methods that have been employed in these successful certification programs so that this information may be used as a basis for evaluating plans and proposals for future demonstrations of design adequacy and acceptability.

Workload evaluation techniques that have been used to assist in satisfying the requirements of the Federal Aviation Regulations are described and catalogued. Known limitations on utility or validity of data are summarized.

The authors wish to express their appreciation to the Boeing Commercial Airplane Company, the Lockheed-California Company and the Douglas Aircraft Company for their willing assistance, data, illustrations, helpful comments and cooperation provided during this documentation effort.

Recognition is also given to the Aircraft Certification, Engineering Flight Test and Flight Operations Evaluation organizations within the Western and Northwest Regions of the Federal Aviation Administration for their assistance and useful comments provided during the course of compiling and reviewing the material for this report.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (CONT'D)

EXECUTIVE SUMMARY

To establish the minimum crew required for safe operations, the FAA evaluates crew workload: (a) with the number of crewmembers dictated by cockpit design; and, (b) in relation to a projected operating environment.

This report describes the techniques and procedures used to evaluate workload successfully in past airworthiness certification programs. It also summarizes related workload assessments made during aircraft design. There is no intent to establish a formula for future crew complement determinations, as it is recognized that workload measurement is a developing science. The FAA is conducting human-factors workshops and is supporting a broad program of study which may accelerate progress in this and related areas. Emerging workload study methods and research programs will be documented in future publications.

Past workload evaluations involved the consideration of workload implications from earlier concept studies and continued as appropriate through various mockup, simulation, analytical, experimental, and flight test procedures. The multifunctional nature of the human operator necessitates the use of a variety of evaluation techniques and procedures. When applied at appropriate stages of development, these methods serve to answer crucial questions about workload. Final evidence of design adequacy is developed in flight test, because neither simulation nor analysis, without actual flight operations, can provide total substantiation that workload and crew duties are satisfactory when compared to existing operational aircraft.

The simulation methods employed to date are most useful for demonstrating overall configurational suitability and specific stimulus-response adequacy. In addition to their use for appraising procedures, task features and resulting workload implications, mockups are used to test the visibility and conspicuity of indicators, the convenience of reach and accessibility of controls, and the conformance to layout conventions and pilot expectations. More functional simulators are used to measure the complexity and number of required procedures by count, and by timing simulated pilot actions. The ease of operation of controls and utility of warnings are among the questions examined, and in some cases comparisons are made between activities using new design features versus features of an existing, service-proven flight deck design.

Despite the great utility of simulation, not all problems can be solved this way; particularly person-to-person interactions in simulation do not duplicate routine flight conditions due to motivational differences--hence many causes of errors cannot be revealed. Also, simulation is not sufficient to prove the operational suitability of large changes in cockpit design, such as conversion to electronic flight instruments. Major changes in the flight crew/systems interface may require a complete cycle of analysis, simulation, and flight test before sufficient understanding of the interface is achieved to permit application to operational transport aircraft.

For example, the electronic flight instruments being used for the new generation of commercial aircraft were first evaluated in simulators during the early 1960's, initially flight tested in the late 1960's and have been flown in FAA-NASA flight tests and demonstrations since 1973 aboard the TCV aft flight deck.

Analyses are made using computer models of internal visibility and physical action requirements. More elaborate time and task computations are also made using pilot response data from earlier detailed part-task recordings and procedures-required tabulations from sample flights in high workload regimes. Comparison data on time to complete actions in the new design in contrast to an operational flight deck are presented to show the balance of workload between crewmembers and the appropriate distribution of work requirements over busy periods, such as approach and landing.

Flight test, including a simulation of routine airline operations, is employed to substantiate the adequacy of design and the acceptability of emergency procedures as well as to demonstrate ordinary flight duties which are characteristic of the new design.

There is no simple solution to all the issues raised and no single tell-all method of testing new designs prior to availability for actual flight. However, workload evaluations can be accomplished to satisfy needs during aircraft design and to provide needed numerical data to support test pilot subjective ratings of acceptability. In combination, the various assessment procedures have been successful. Aircraft designed to be flown by different crew complements have been so certificated and have been proven safe and acceptable in actual line service. The correlation between FAA workload determinations in certification procedures and the ultimate criterion of airline experience has been excellent. Still, the many difficult decisions made in designing and approving complex certification programs should be recognized, and research efforts should be extended to develop improved test methods as the understanding of human behavior allows.

Since a portion of all successful evaluations currently involves the attitudes and perceptions of flight crewmembers, there will be a continuing need for subjective assessment. It is essential that these assessments be made by persons who are experienced in conducting procedures in differently designed cockpits and who are accountable for their judgments. Otherwise, strong individual bias may influence pilot opinion.

In summary, workload confirmation is a continuing process from the earliest concept development through the successive design and development stages. Ultimately, confirmation is accomplished in the prototype airplane as it undergoes intensive test and evaluation scrutiny to confirm and demonstrate suitability. Also, it is noted that individual production aircraft are examined to consider the workload impact of equipment or configuration variations. Finally, each airline is inspected to verify that actual flight operations are satisfactory with the unique combination of pilot qualifications, special airline procedures, flight deck equipment outage allowances, company equipment added, and challenges of the particular operating environment.

1.0 Introduction

1.1 Problem

Each transport aircraft must have a Type Certificate as the basis for its production, and an airworthiness certificate issued by the Federal Aviation Administration (FAA) before it can be operated in regular airline service. Before the award of the Type Certificate, the FAA conducts tests and evaluates data submitted by the manufacturer, who is the applicant, to determine compliance with design criteria and airworthiness sections of the Federal Aviation Regulations (FAR), Part 25, Airworthiness Standards: Transport Category Airplanes (See references 1 and 2).

Assessment of crewmember workload is an element of new aircraft certification as provided in FAR 25.1523, title "Minimum Flight Crew," and Appendix D. In addition to section 25.1523, other provisions of Part 25 state various flight deck design requirements, with particular sections having to do with instruments, displays, controls, and miscellaneous equipment such as seat restraints, electric protective devices, radios, fire extinguishers, etc. In the type certification process, it is the application of the minimum flight crew section, 25.1523, that provides the most general check on the adequacy of human factors design in the flight deck. In turn the workload factors enumerated in Appendix D, will be seen to constitute a comprehensive human engineering checklist including: the accessibility, ease and simplicity of operation of all necessary flight, power, and equipment controls; the accessibility and conspicuity of all necessary instruments and failure warning devices; the number, urgency, and complexity of operating procedures; the degree and duration of mental and physical effort; the extent of required monitoring functions; the actions requiring a crewmember to leave assigned duty station; the degree of systems automation; the communication and navigation workload; increased workload in emergencies and with flight crewmember incapacitation. Workload assessment serves, then, as the context in which overall flight deck design is evaluated and approved.

The immediate problem arises not from any discrepancy between past workload determinations and the final proof of acceptability of crew size in actual airline operations. In fact, the crew size determinations made during original type certification proceedings have been reviewed many times in the light of later aircraft service experience, and it has never been found necessary to change the crew size specified in an original turbojet airworthiness certificate. A 10-year review of accident data made by the FAA Task Force on Crew Workload resulted in the conclusion that no safety problem exists with the current operation of 2-crew air carrier aircraft (see reference 3). An update of the accident summary covering the years 1977, 1978, and 1979 confirmed the fact that 2-crew transport aircraft are being operated safely. Furthermore, an appended study of rules violation showed no higher incidence of violations has occurred with the smaller size crew (see reference 4). Still, it appears desirable at this time to provide more information on the general subject of flight-deck design approval and just how the decision is made that a particular flight crew composition is safe and appropriate for a given aircraft (see reference 5). Many of the statements made in published articles by protagonists of

the view that all large transports should have three pilots clearly show a lack of understanding regarding the influence of design on procedures and tasks, and thus on workload and the kind and magnitude of testing that is conducted prior to the award of a type certificate. This part of the problem may be corrected by providing an account of the flight deck engineering development methods employed in previous successful certification programs.

Simply stated, there is no problem with the crew complements of transport aircraft. Throughout the modern era some aircraft have been designed for and operated by two flight crewmembers, others with three, and both crew sizes have been entirely adequate in safety and efficiency.

1.2 Objective

The objective of this report is to describe the crew complement determination programs that have been accomplished under the present regulations in the more than sixteen years that those rules have been in force. This description may be most useful if it is organized in such a way that persons with a particular interest in experimental evaluations and other workload assessment techniques that have been employed at various stages of aircraft programs may find those procedures described by program stage: commencing with predesign trade-off studies, during design, after design, and during flight test. Since multiple applications are made of many workload techniques, this plan makes a degree of redundancy inevitable.

In addition to repeating similar descriptions of applications of workload procedures at different stages of aircraft programs, a degree of complexity is added to this subject due to the fact that many different kinds of tests and evaluations are carried out. The necessity of this stems from the many facets of human capability and behavior which are involved in the task of operating a transport aircraft. The final conclusive proof of flight deck design adequacy and acceptability of workload is obtained after the fact through the experience of years of actual line service for the aircraft. While flight tests conducted before certification of the aircraft have always been successful in assessing the acceptability of the design crew complement, as judged by this final criterion, the same does not necessarily hold true for various individual simulation and part-task test techniques that are frequently employed prior to the availability of a finished test aircraft. For this reason, it is not desirable to select just the one best or the few most productive design stage test methods and apply such a reduced inventory of workload procedures to all flight deck adequacy questions. Instead of this, a wide variety of test and evaluation techniques are adapted to varied and specific questions that arise during the design and later operational suitability test stages of an aircraft program. Different procedures have been found to be most powerful in developing data on different questions, but all methods short of full flight test have particular best applications and particular limitations that need to be recognized. Clearly, the surest way to demonstrate acceptability of workload demands is to fly the aircraft. But since this cannot be done earlier in the aircraft program while potential customer requirements are being converted into cockpit design specifications (from 3 to 6 years before flight), reliance must be placed on test methods that are less conclusive and less exact. It is the variety and utility of these pre-flight test procedures that becomes the primary subject of this document.

Over a period of time much has been learned. Workload evaluations methods have evolved in parallel with the creation of successively more automated aircraft systems and in step with the understanding of the basis for human performance and perception. Initial confirmation that procedures and tasks are in fact feasible is based on analytic methods, sketches and drawing reviews. Flight deck mockups have always been used with specialist pilots employed to evaluate physical arrangements and crew procedures. Today the "engineering" simulator has come into use to permit more dynamic testing and more flexible alteration of components, configurations, and flight equations. Similarly, the conduct of mock procedures in a ground cockpit has been a time-honored method, but recently video and computer developments have made it possible to generate more data and to include more diverse test conditions. One goal of this report is to show the present state of this technology by describing tests conducted in recent certification programs.

1.3 Critical Issues

There is, of course, an orchestrated effort by pilot associations to establish one, single, crew configuration for future turbojet transport aircraft. The flight deck design favored by these groups requires three crewmembers, a pilot (Captain), a co-pilot (First Officer), and a side-facing third person variously designated a flight engineer or third pilot (Second Officer). This configuration is similar to that of the majority of current turbojet transports, with the B-727 the most numerous and widely used example. However, about one-fourth of the domestic airline turbojet fleet, accounting for about one-third of all departure cycles, use an alternative cockpit designed for operation by two crewmembers.

In the development of the arguments, several distinct issues have been raised by those defending the case for making the B-727 type crew configuration an industry standard, by the airlines and aircraft manufacturers who have in many instances opposed that plan, and by the FAA in attempting to clarify the federal role established by law with its attendant responsibility to all parties. In this section, five of the most significant issues related to flight deck workload evaluation and certification will be defined. These issues are: airworthiness criteria, standard crew size, validity of simulation, FAA and the applicant, and certification participation.

1.3.1 Airworthiness Criteria

It is the whole aircraft that is examined and determined to be airworthy, not component aircraft parts such as the crew compartment and its associated systems. Still, the importance of flight deck design, and the ability of the planned crew complement to operate in that design environment, to overall aircraft performance is so great that evaluation of a new flight deck becomes a critical part of an overall aircraft airworthiness determination. What standards are appropriate, then, to be used as criteria of airworthiness?

The first criterion is aircraft performance in the prospective operating environment. It must be determined that the design and crew are capable of safe operations in the airports, enroute environments, ATC systems, and the ancillary operating contexts such as navigation-communication provisions and maintenance systems.

The second criterion is acceptance of the required procedures and resulting workload. This means that an aircraft found to be excessively difficult to operate or overly fatiguing will not be approved. An interaction between the two criteria may be recognized in that a pilot might be expected to rate workload excessive if he were not able to perform all the required procedures to ensure safe operation. In that sense, the two criteria merge into one standard--the aircraft compares favorably with other proven, inservice aircraft in operability in the anticipated environment and in demands and provocations affecting the crew.

Some have advocated a requirement for three crewmembers on all future aircraft in the large transport category and say that the correct criterion of airworthiness should be "optimum" design or a more than acceptable and comparable rating (see reference 5). The question of the legal rights of manufacturers to equitable treatment by the FAA will be discussed in a separate issue below. At this point, it should be noted simply that it is not required that a third engine be installed on the theory that three give a greater margin of safety against mechanical failure than two, and, similarly, it is not required that each passenger have a custom fitted seat restraint system just because such devices would add to survivability in the case of an accident.

The cornerstone concept is that current airline operations are safe. This is proven by comparison with either past records or competing transportation modalities. If everything else were equal, adding an extra engine or some cumbersome "safety" device might make airline operations even safer, but everything else is seldom equal. Money spent, airframe allocated, and lift devoted to carry an unneeded propulsion unit would not be available for other uses that may be more important. This logic is applied to all aircraft systems, so that the rule is that a new aircraft is airworthy if it conforms to established requirements and compares favorably with other accepted and service-proven aircraft. A requirement that the unprovable, or attainment of "optimum' safety or efficiency, be demonstrated is inconsistent with the philosophy which has regulated the air transportation industry as it developed into a safe, dependable service.

1.3.2 Standard Crew Size

Two-pilot crews have co-existed alongside three-man crews for many years, just as two engine aircraft have operated in the same environment with other propulsion configurations. No number of engines is standard, and no number of flight crewmembers is standard. In addition to the variation in number of crewmembers, there are two main layout variations in three-man crews. In all cases, the captain, or pilot-in-command, and the co-pilot, or first officer, are seated side-by-side, forward facing. If a third crewmember, or second officer, is provided, there may be a system display and control panel placed outside the forward view of the captain and first officer. One version of this layout, called the sideward facing cockpit configuration (SFCC) seats the third crewmember behind the co-pilot and facing sideward toward his own panel. A swivel seat allows the second officer to turn and face forward when desired, but most of the time he is removed from the peripheral view of the forward facing pilots. An alternative layout for three crewmembers is called the forward facing cockpit configuration (FFCC). In the FFCC, the second officer sits between, and

slightly behind, the two forward pilots. From this vantage, the third crewmember shares the front panel, pedestal, and overhead panel presentations with the other crew members.

Before the changeover to turbine engines, pilots had several essential flight responsibilities that have either disappeared completely or have been reduced in prominence. An obvious example is that powerplant controls have been simplified to power levels with the elimination of mixture, and in turbojets, the elimination of propellers. Similarly, radios have been simplified with digital tuning. Generally speaking, fuel transfer systems have been simplified, and in some recent cases, eliminated.

Many other items could be listed, but the point is clear -- modern transport aircraft are designed and built to be more reliable and make fewer manual adjustment demands and fewer monitoring demands on the flight crew. The earlier piston engine transport aircraft in which provision was made for a third flight crewmember were a minority of the piston types, despite the more complex required activities. Most early transports had only two pilots. Any additional crewmembers were carried as navigators, radio operators, or mechanics.

When a third flight crewmember was added in what came to be known as a side facing, flight engineer's station, the added crewmember was provided because specific required tasks were difficult for the forward facing pilots to accomplish. For example, some fuel systems necessitated fuel transfer from tank to tank which was accomplished by opening and closing valves with the general character of a plumbing system. A few aircraft had complex engines with many separate ignition elements that were unreliable and had to be monitored. A full panel of small but important indicators was required to do this, and that panel could not be located conveniently for observation by forward facing pilots. Reasons in this category underlay the initial requirement for a third flight crewmember in a minority of the more complex early transports. Still, as stated above, there was not a standard crew size established -- most piston transports carried two pilots.

The first generation turbojets also showed a diversity of crew complements but coming under the then current FAA rules, size alone mandated three crewmembers in the widely used four-engine transport aircraft, the B-707 and DC-8. Subsequently designed turbojet transport airplanes either used the cockpit design of prior models to save design, certification and training costs, or, in smaller models, used cockpits designed specifically for a two-crew operation. This has tended to imply the existence of a convention that relates the smaller, two-engine airplane with a crew of two and the larger three- or four-engine airplane with a crew of three. Opposition by pilot associations was voiced to the two pilot configuration in smaller turbojet transports, but only in isolated instances of individual airline labor-management agreements were the unions successful in obtaining a position for a third crewmember. Hence, at the present time, about one-fourth of the U.S. domestic airline turbojet fleet operates with two crewmembers while the remainder operate with three.

Airline flight decks vary, then, not only in the number of required crew members but also in the configuration in which the crewmembers are seated. During the controversy over plans for future aircraft, proponents of

three-man crews have criticized the FFCC as being susceptible to operation by two, as well as by three, crewmembers. Apparently, one reason that the SFCC is preferred is that it removes system displays and controls from the reach of the forward crewmembers and makes normal operation with less than three impossible. Otherwise, various opinions have been expressed as to which crew complement and which alternate seating arrangement is most "integrated" (see references 6 & 7). The one agreed certainty is that current production turbojet transports vary in crew provisions and do not exhibit a standard crew size.

1.3.3 The Validity of Simulation

The advocates of a rule requiring future turbojet transport aircraft to provide for three flight crewmembers have stated that the simulation state of-the-art makes it feasible to compare alternate two- and three-man designs. All transport pilots are familiar with modern flight simulators, since airlines use simulators for recurrent and transition pilot training. Most of these "training" level simulators have moving bases to provide cues to aircraft accelerations and have high quality visual attachments to permit practice of transition from instrument flight to visual contact and the execution of approach and landing maneuvers. The FAA has approved the training simulators for substitution in place of practice of most maneuvers and emergency procedures in the actual aircraft, and the improved safety records of current airlines more than support that substitution.

In the past, a number of fatal accidents resulted from inflight practice of engine failure and other system failure procedures. Today, most authorities believe that simulators offer a better environment in which to experience provocations of this kind and to practice correct problem solving. Then, if the pilot ever encounters similar problems in flight, he will be skilled and confident in his response. Hence, the validity of training level, or full performance simulators for pilot instruction, practice, and crew coordination training is generally accepted.

It is a considerable step to generalize from the training validity of simulators to the conclusion that a new flight deck design, for which there is no inflight experience, can be properly assessed as to total pilot workload from these devices. There are reasons to believe, as discussed below, that even the comparison of two versions of a flight deck design is a high-risk undertaking.

Simulation has been employed at various levels and at several stages of flight deck design and assessment. Later sections of this report will summarize simulation techniques ranging from mockup procedures studies, to part-task activity analyses, to computer replicated task-time studies. It will be seen that various forms of these simulation techniques have been successfully employed during the flight-deck design process and also in workload assessment procedures after the completion of that design. Comparison between the new feature, subsystem, part configuration, or overall integration on the one hand, and a comparable segment of the flight deck of a comparison aircraft on the other, has also been employed successfully in the past. Hence, it should be recognized that simulation

is one of the valuable and valid techniques of workload evaluation. At the same time the point is made that simulation is a valuable and even a necessary part of the overall evaluation process, it is presently believed that simulation alone is not sufficient to determine airworthiness, as defined in issue 1.3.1 above.

The flight simulators that are ultimately used for approved crew training include systems that have been flight tested and employ display-control responses that have been derived from and compared to those of the actual aircraft and found to be correct, or at least adequate for training. Thus, the training level simulator is built in final form after flight test of the aircraft, not before, and the flight equations as verified in actual flight test are incorporated in the dynamic computations where necessary to ensure transfer of training. Workload tests conducted in a full performance simulator not so validated, because they were being conducted before the actual aircraft was built and flight tested, would not have the same probative value.

A second problem found in earlier studies is that the result of a new procedure or system in a simulator may be different from the result in an air craft. The reason that this disparity of result may occur is related to the unknown differences between properties of the pilot stimuli in the simulator and aircraft. For example, a more complex situation display may be found to be acceptable and helpful in flight operations while a simpler display lacking some elements of computed information may be preferred in the simulator. This might occur because, in the air, the pilot has additional cues to interpretation that permit integration of more information. How to compensate for such sources of error is unknown.

Finally, there is a severe methodological problem in comparing operations of a two-person crew in a new aircraft design with operations of a three-person crew in a parallel design. In a test conducted in a simulator, it is expected that each crewmember will carry out the standard behaviors that have been instilled in training. Boredom will not cause inattention because of the added drive provided by the test situation. Irrelevant conversation will not occur because of the presence of the simulator operators or test coordinators. Shortcuts, omissions, or other divergent activities will not occur for the same reasons. How, then, can generalizations be made to the real world where it is known that adherence to required procedures and increased cockpit discipline would be important aids? Compounding the method problem is the question of what comparison should be made. Should it be two versus three crewmembers conducting the same procedures in the same cockpit with redundancy added by the third person; or should it be two in a design requiring two versus three in a design with less automation and system simplification, so that there exists a group of genuine and essential duties exclusive to the third crewmember?

The specific recommendation recently made by pilot associations was that a full-mission simulation program should be conducted to compare alternative cockpit configurations of advanced transport aircraft. Taken as a research proposal, and with the proviso that a broad range of technical developments including increased system simplification and flight deck automation be incuded in the comparison, this appears to be a potentially fruitful activity. It has, however, for the reasons outlined above, no relation to

the certification of any one aircraft under FAR 25.1523. It is not the case that a simulator or family of simulators can be built to fully duplicate actual inflight workload, and an approximation of that fidelity is not verifiable prior to flight test of a given new aircraft. Alternate concepts of flight deck systems, management procedures, and display-control relationships can be studied and compared in simulation, and in flight. Such projects can add to the overall store of knowledge that is used to guide the design of future aircraft. Taken in this context, the current proposal is not new and does not cover additional activities beyond those envisaged in development studies already underway at government and industry laboratories. What was new was the proposal to substitute the results of a simulator comparison of different versions of a new aircraft flight deck for the usual study procedure, which is to compare one prototype configuration for the new aircraft with a single flight deck that has been proven acceptable in line service.

Comparing one new flight deck design with another, both unproven in service, cannot provide the same degree of assurance that is obtained in a program that compares a new design with a proven design. This latter procedure need not make the same assumptions about the correlation of the total simulation test environment with the real world, because attention may be focused on those individual flight deck features that have been changed. If it is shown that pilots perform better and are more approving of a new display, a control, or a pattern of elements than they are of a standard, service-proven corresponding element, a substantial point of information has been acquired. We know that the service proven element is acceptable, whether a full simulation including that element is truly representative of real-world workload or not. Hence, finding that the new element is equal to or better than the standard, again whether the entire test situation is fully representative or not, is a highly valuable test result.

Two other proposals made for acquisition of data related to airworthiness certification are that the crew survey technique be used and that the FAA obtain independent verification of manufacturers' failure rate analyses. At this time, there is no definite plan to include these two areas in the FAA human factors program for the following reasons.

The operational crew survey technique is not new. It has been used in the past and found to be highly unreliable and subject to individual bias. A given crew is usually employed in operating a particular aircraft type and is not intimately familiar with other designs and procedures. To a great extent, line crews are highly oriented in present approved practices, and this orientation makes it hard for them to imagine the impact of changed flight deck features and procedures. As a result, the problems they identify and the estimates they make of the impact of changing crew duties and redistributing crew workload are influenced heavily by their current qualifications and experience in older aircraft. To provide information that can guide future design efforts, and in particular, information that identifies problems in a way that evaluates alternative solution possibilities, a pilot or other evaluator needs a broad perspective. This perspective is generally developed over a long period of time through participation in the evaluation of cockpit systems, crew procedures, atmospheric

environmental effects, flight handling qualities, etc. Credibility of the information provided by cockpit design evaluators is also affected by the degree or level of comment responsibility assumed by or imposed on the evaluator. The more the evaluator becomes responsible for the evaluations and comments offered, the greater the reliability and quality of the evaluations offered. The experienced test pilots employed by the manufacturers and the government and the check pilots working for the airlines have developed many of these characteristics out of necessity. If not developed, either through training or experience they do not remain in this work very long. In contrast, line crews uninitiated into this philosophy and discipline may be limited in their ability to provide suitable insight into the future application of new systems and newly applied technology.

The failure rate analysis question should be seen in full perspective. It is not the case that an avionics supplier, for example, markets his device and receives uncritical acceptance of reliability data. The Airlines Electronic Engineering Committee establishes standards for system elements and criteria for reliability testing. The airframe manufacturer conducts tests on components and systems before selecting suppliers for new flight deck elements, and the customer airlines evaluate the resulting data on a committee basis. There are, then, several stages of system reliability evaluation before type certification proceedings, and the type certification board is empowered by law to make any additional tests deemed to be needed. Furthermore, no data are in hand showing that current failure evaluations are not satisfactory.

If there is substantial reason to believe that critical aircraft systems are exhibiting reliability not congruent with published analyses, the actual record should be studied. Present and feasible future airborne recorders can identify failures and resulting effects on aircraft performance and crew activities. Rather than examining such data through a crew recall filter of possible extreme subjectivity, it seems clear that actual performance data should be utilized.

One proposal that may avoid, at least in part, the limitations of simulation is to follow a single crew in the simulator for a series of "flights" duplicating a crew schedule of at least a month. If this were done, it is possible that the pilots would adapt to the test situation and begin to act "naturally." No data confirming such adaptation is presently available, and it may be noted that taking a group of crews through a month of repeat simulator sessions would be both costly and productive of a large mass of possibly ambiguous data.

1.3.4 FAA and the Applicant

Before the 1980 decision to establish a single "lead region," (i.e., the Northwest Regional Office located in Seattle, as the single FAA regional office conducting transport category certifications) certification testing was accomplished under the supervision of the regional office in whose area the manufacturer was located.

Much of the routine work of engineering inspection and documentation has been delegated to Designated Engineering Representatives (DER's) who are in the employment of the certification applicant. The FAA, having direct access to the DER's without the applicant's review or approval, uses them

to review and evaluate a limited range of structures or systems. The purpose of the DER system is, of course, to stretch the capability of the relatively small FAA engineering and manufacturing staff so tnat all important operations may be monitored. These two facts, the close, long term proximity of FAA regional office to manufacturer, and the delegation of limited approval authority to DER's, have been cited by critics of the FAA as indicating a lessening of the desired degree of independent government review and quality control. It is the case, however, that, in the type certification procedures, FAA regional offices have never designated a DER for a transport aircraft flight deck, as they have for passenger compartments, for example. The institution of the lead region includes provision for a cadre of FAA certification specialists who will be assigned to that region for the critical period of a type certification proceeding. Hence, experts in matters such as crew workload will be brought into the Northwest team specially for the upcoming programs, and in general, these temporary additions to the regional staff will not have experienced the day-to-day contact with the applicant during previous design and test phases of the new aircraft program. This tends to ensure that an independent review of workload data will be made without possible carryover effects from previous close contact with the evolution of unique design features.

In the final analysis, no practical way is known to isolate, completely, FAA certification authorities from knowledge of previous tests conducted by the manufacturer and previous decisions as to flight deck systems by prospective customers, nor would such complete isolation necessarily be desirable. The persons reviewing certification data are going to know that prior to certification, many and varied tests have been conducted successfully and that flight operations specialists from customers have reviewed, approved, and probably paricipated in the new flight deck design. An assertion has been made that FAA is too closely associated with the manufacturer and that this reduces the validity of the certification process. Alternatively, that same close familiarity with program evolution and prior testing can be viewed as a strength, when it is recognized that, to a substantial degree, both the applicant and the FAA have the same goal: an aircraft that compares favorably with previous and alternative aircraft. If, as has been implied by some critics, the manufacturer sought to improve marketability by reducing safety margins in flight-deck design or other aircraft feature, and succeeded in obtaining FAA approval of the substandard aircraft, that would constitute a highly counter-prcductive course of action. To make an aircraft program economically successful, usually additional aircraft must be sold for more than ten years after the initial deliveries. A substandard aircraft would be unmasked before that time had elapsed, guaranteeing a losing program for the applicant.

It may be fair to say that in the conduct of life-and-death regulatory matters, such as airworthiness certification, some reasonable degree of moral rectitude should be assumed of government functionaries. The FAA is a continuing organization in which people are accountable for their actions. In the absence of a historical incident of improper influence by a certification applicant, and on the record of world leadership in transportation safety progress, it is unreasonable to advance vague charges that tend to impugn the independence of those responsible for certification.

A final aspect of the relation between the FAA and the applicant is the matter of the rights of the applicant. A manufacturer is free to expend his private resources in the design and development of a planned product. The law establishing the FAA assigns to the FAA the authority to certify a new aircraft or to withhold certification, in which case the aircraft could not be sold and the manufacturer would lose the investment in design and development.

The due process clause in the Constitution, the Administrative Procedure Act, associated legislation, and precedent make it clear that the FAA may not refuse certification except on good and sufficient grounds. If a new aircraft is equivalent, or better than, competing products, and there is adequate evidence that it can be operated in the national airspace system with safety and no undue interference with that system, such as violation of noise limits, the FAA has no authority to withhold certification, nor would anything be gained by such action.

Much of the criticism advanced on workload evaluation includes the demand that proof be given that the applicant has optimized the crew facilities and procedures. Leaving aside the question of what constitutes an optimum pilot workload, although recognizing this to be a complicated subject in itself, there is no basis in law for the FAA to require such optimization. To do so would be to deprive an applicant of his right to equitable treatment vis-a-vis other manufacturers.

The actual language of the act creating the FAA, Public Law 85-726, should be noted. "General Safety Powers and Duties" are enumerated in Title VI. Under the subheading "Minimum Standards; Rules and Regulations," the Administrator is empowered" ...to promote safety of flight ...by prescribing ...(1). Such minimum standards governing the design, materials, workmanship, construction, and performance of aircraft, aircraft engines, and propellers as may be required in the interest of safety..." Under the subheading "Aircraft Certificates," the act continues to state that: "The Administrator shall make, or require the applicant to make, such tests during manufacture and upon completion as the Administrator deems reasonably necessary in the interest of safety..." Further, the act states; "If the Administrator finds that such aircraft, aircraft engine, propeller, or appliance is of proper design, material, specification, construction, and performance for safe operation, and meets the minimum standards, rules, and regulations prescribed by the Administrator, he shall issue a type certificate therefor " (see reference 1).

1.3.5 Certification Participation

It has been asserted that pilot associations should be parties to the airworthiness certification process, should be allowed to examine test data prior to certification, and should have a right to present other facts and opinions to be considered by the FAA before certification. Under the law, only FAA and the applicant are parties, and the FAA is required to protect the proprietary information that is provided by the applicant in support of the request for certification. Hence, requests for a direct role in the certification process made by others have been rejected, although all interested persons have been encouraged to offer any information, suggestions, or other comments to the FAA with the assurance that all inputs are considered.

In this era of spirited "public interest" groups, one can easily imagine that there are other people who would like to be accorded the opportunity to examine and refute the data collected by the FAA and supplied by the manufacturer. Added to these would be, no doubt, competing manufacturers, including foreign corporations, and overseas aviation regulatory authorities, who, of course, are often closely affiliated with national manufacturers competing with U.S. aerospace companies. The protection of the private property rights of the manufacturer must be considered. This makes the opening of the certification files an extremely sensitive issue.

The airlines who procure the new transport aircraft normally have contracts with flight crewmember associations that have been formulated on the basis of labor-management negotiations. The pay, working conditions, crew scheduling procedures, and the like applying to pilots and flight engineers are specified in such contracts. Through negotiations, different airlines have reached different agreements on all of these provisions with their separate crewmember associations. In at least two cases, such resulting contracts have required the employment of a third flight crewmember in an aircraft certificated for a crew of two and flown by many other airlines with that crew. Traditionally, the accepted view has been that staffing is a management responsibility subject to agreement with labor. The government does not set requirements for the number of persons who must be employed to perform jobs except in the case of safety regulation, such as the determination of the minimum flight crew for transport aircraft. Operators are normally free to exceed such minimum staff requirements as determined in agreements with labor.

There is a problem with optional aircraft crew-size determination by labor-management agreement that has only recently come to be recognized. This is that the data examined during certification to ensure that the aircraft is safe and can be operated appropriately in the national system are all collected with the design crew size, and additional information on the performance of alternate crew sizes is not analyzed. Also, the typical seat placement preferred by the additional crewmember and the assigned duties of that crewmember may conflict with other requirements, such as the provision of an operations inspector's position. When an FAA inspector or a company check pilot is performing his functions, the additional crewmember must be removed to a rear seat usually out of reach of essential information and controls. This alternate seating breaks down the standard crew duties and standard procedures that are believed to be important to safety and generally tends toward variability of performance. From these considerations, it can be deduced that the addition of crewmembers, the reassignment of essential flight duties, and the deviation from the crew performance standards that were enforced during certification testing greatly reduces the assurance that the airworthiness process is intended to provide.

The individual airline is inspected by FAA district office personnel, and its special operating procedures are reviewed and approved for safety. While this individual airline inspection and regulation procedure has functioned satisfactorily in the past, there is concern that increases in automation and the rapidly evolving digital instrumentation and flight management systems may dictate change in future requirements. It is the lessened tolerance of integrated, sophisticated systems to individual behavioral variations that calls into question the validity of allowing wide

deviations in crew provisions. If the aircraft control systems were demonstrated, tested, and approved for one standard crew size, minimum equipment list (MEL), and specific crew performance objectives (CPO), does it follow that line check inspections can be expected to provide adequate data on which to base approval of airline operations that deviate from the conditions of original testing and certification approval? This question assumes a new urgency in the case of the forthcoming Boeing aircraft designated the B-757 and B-767. It is understood that both aircraft will incorporate major automation and digital system advances and that there may be alternate versions of the basic flight deck design (see reference 8). If one aircraft has a cockpit configured for three pilots, is tested, and certificated for operation by three, it will not be legal to fly that aircraft with less than three crewmembers. Similarly, if another aircraft version is configured for a flight crew of two and is tested and certificated under those conditions, should the option be left open to revise crew procedures and operate with three? It appears to some authorities that the workload evaluation techniques found to be appropriate in any given program can only yield results applicable to the conditions of the tests. Adding to crew size, revising standard equipment performances, or deviating from CPO's may invalidate the process.

These considerations make it necessary to re-examine the issue of the latitude allowable for labor-management negotiation. It seems possible that a full examination of this issue and its ramifications for safety will lead to the conclusion that the crew size established in certification must be adopted as the operating crew size for that aircraft industry-wide.

1.4 Approach

1.4.1 Body of this Report

The remainder of this report consists of a chapter addressed to fundamentals, three chapters that detail workload evaluation techniques employed in past programs at different chronological stages of aircraft development, and a chapter on documentation and reports. Under the heading of fundamentals, the regulatory history of crew complement determination is reviewed, and procedures appropriate to workload assessments during aircraft design and after aircraft design are introduced. The following three chapters take up these procedures in more detail, with illustrations drawn from particular past certification programs. Discussion of workload evaluation techniques that are now being proposed for use in future programs will be treated in a future report. The present report is historical in nature.

1.4.2 Type Certification vs Aircraft Certification

The focus of this report is on aircraft type certification. It is recognized that, in fact, there is an airworthiness certificate issued for each individual aircraft manufactured in accordance with the type certificate itself. Different airline customers may, and typically do, order variations in new aircraft cockpit equipment. When such an airline request relates to flight deck features, the manufacturer examines the workload impact of the difference and discusses any identified impact with the FAA.

When the specification becomes firm, a "Letter of Definition" is prepared outlining the major systems and characteristics of the ordered aircraft, with significant variations described. Typically, the changes will not be considered crucial, and the manufacturer will recommend that the individual aircraft be approved by a Designated Engineering Representative (DER) at the company. Sometimes, it will be requested that the FAA determine the eligibility for certification of the individual aircraft. Also, it happens on occasions that the FAA does not agree with the proposal to have approval decided by the DER. In these cases, FAA flight test personnel examine the aircraft directly. An example of correspondence relating to this process is shown in Appendix A to this report.

The full regulatory context, then, consists of four essential stages. First, a specific requirement to obtain a type certificate is that the crew complement must be determined. Second, the manufacturer must obtain or have a production certificate which assures that the manufacturing and quality control procedures provide conformity of the airplane with the approved design. Third, individual aircraft are approved and issued airworthiness certificates after consideration of any changes made since the type certificate proceedings. Fourth, the operating airline is inspected; its training, MEL, required procedures, maintenance practices, etc., are reviewed and subjected to continuing scrutiny by inspectors located at nearby district offices.

Under the present system, as outlined above, there is a continuing learning process. While the basic characteristics and minimum operating limits of the aircraft are defined during aircraft development and type certification, changes are incorporated in standard equipment during the production life of the aircraft type, and individual customer variations are introduced. Continuing monitoring by the inspectors assigned to each airline contributes to the growth of information related to each aircraft type. Thus, when the B-737 was certificated for operation by a flight crew of two, there was no operating experience other than that gained with initial production aircraft during type certification. More than a decade later, there is available a wealth of information covering many versions of the basic aircraft equipment list, many airline operating procedure variations, and safety statistics covering widely varying flight environments.

Accrual of information over years of airline operations enhances the confidence with which decisions concerning the best operating procedures can be made. Regulatory authorities, like manufacturers and operators, benefit from an experience learning curve effect. One aspect of this learning curve is applicable to crew complement determination in type certification. That is that world-wide airline experience with such two crewmember turbojet transports as the BAC 1-11, the DC-9, and the B-737 has been successful (see reference 9). This success adds assurance that the workload evaluation procedures and techniques employed in past programs have themselves been successful in predicting the final, real-world results. This past success makes it all the more worthwhile to record those procedures and techniques so that all interested persons may become familiar with the present available pool of techniques.

2.0 Fundamentals

2.1 Certification Requirements for Minimum Flight Crew

Certification of transport category airplanes is done by showing compliance with the applicable airworthiness standards. This is done for the design of power plants, equipment, structure, and flight performance, as well as for the minimum crew requirements.

Federal regulations adopted in 1948 required that air carrier aircraft with a takeoff gross weight in excess of 80,000 pounds must utilize a third flight crewmember. The 80,000 pound figure was established at that time to exclude the DC-4 from the third crewmember requirement while including the DC-6 and subsequently certificated heavy aircraft. As a result of this regulatory requirement, the over-100,000 pound gross weight Caravelle was required to have three flight crewmembers while the BAC-1-11, being slightly under the 80,000 pound criterion, was certificated with only two crewmembers.

In June, 1960, an FAA conference was held to determine if the 80,000 pound rule should be replaced with more realistic criteria. After considerable research and evaluation, in January, 1964, the 80,000 pound rule was revoked by the FAA, with notice of revocation issued April 27, 1965, retroactive to January 1, 1964.

In November, 1964, the FAA recodified the new airworthiness standard for transport category airplanes with an effective date of February 1, 1965. The new standard, Federal Aviation Regulation (FAR) Part 25, replaced the Civil Aeronautics Regulation (CAR) 4b.

The revision of the old CAR 4b rule after the deletion of the 80,000 pound rule states:

> "Minimum Flight Crew - The minimum flight crew shall be established by the Administrator as that number of persons which he finds necessary for safety in the operations authorized under Section 4b.721. This finding shall be based upon the workload imposed upon individual crew members with due consideration given to the accessibility and the ease of operation of all necessary controls by the appropriate crew members."

The revision to the new FAR part 25 states:

> "The minimum flight crew must be established so that it is sufficient for safe operation, considering --
>
> (a) The workload on individual crewmembers;
>
> (b) The accessibility and ease of operation of necessary controls by appropriate crewmember; and,
>
> (c) The kind of operation authorized under 25.1525.

The criteria used in making the determinations required by this section are set forth in Appendix D."

Therefore, due to the timing of the deletion of the 80,000 pound rule and the adoption of a new FAR Part 25 airworthiness standard, the BAC-1-11 was certificated under CAR 4b under the 80,000 pound criterion, the DC-9 was certificated under CAR 4b but after the 80,000 pound rule was rescinded, whereas the B-737 was certificated under the more stringent Part 25 requirements.

The criteria used as shown in FAR Part 25, Appendix D (See below) list six functions and ten factors to be considered in determining the minimum flight crew.

PART 25 Airworthiness Standards: Transport Category Airplanes

APPENDIX D

Criteria for determining minimum flight crew. The following are considered by the Agency in determining the minimum flight crew under 25.1523:

a. Basic workload functions. The following basic workload functions are considered:
(1) Flight path control.
(2) Collision avoidance.
(3) Navigation.
(4) Communications.
(5) Operation and monitoring of aircraft engines and systems.
(6) Command decisions.

b. Workload factors. The following workload factors are considered significant when analyzing and demonstrating workload for minimum flight crew determination:
(1) The accessibility, ease, and simplicity of operation of all necessary flight, power, and equipment controls, including emergency fuel shutoff valves, electrical controls, electronic controls, pressurization system controls and engine controls.

(2) The accessibility and conspicuity of all necessary instruments and failure warning devices such as fire warning, electrical system malfunction, and other failure or caution indicators. The extent to which such instruments or devices direct the proper corrective action is also considered.

(3) The number, urgency, and complexity of operating procedures with particular consideration given to the specific fuel management schedule imposed by center of gravity, structural or other considerations of an airworthiness nature, and to the ability of each engine to operate at all times from a single tank or source which is automatically replenished if fuel is also stored in other tanks.

(4) The degree and duration of concentrated mental and physical effort involved in normal operation and in diagnosing and coping with malfunctions and emergencies.

(5) The extent of required monitoring of fuel, hydraulic pressurization, electrical, electronic, deicing, and other systems while enroute.

(6) The actions requiring a crewmember to be unavailable at his assigned duty station, including: observation of systems, emergency operation of any control, and emergencies in any compartment.

(7) The degree of automation provided in the aircraft systems to afford (after failures or malfunctions) automatic crossover or isolation of difficulties to minimize the need for flight crew to guard against loss of hydraulic or electric power to flight controls or to other essential systems.

(8) The communications and navigation workload.

(9) The possibility of increased workload associated with any emergency that may lead to other emergencies.

(10) Incapacitation of a flight crewmember whenever the applicable operating rule requires a minimum flight crew of at least two pilots

(c) Kind of operation authorized. The determination of the kind of operation authorized requires consideration of the operating rules under which the airplane will be operated. Unless the applicant desires approval for a more limited kind of operation, it is assumed that each airplane certificated under this Part will operate under IFR conditions (see reference 2).

2.2 Documentation Plans (Measures and Criteria)

The total task that a flight crewmember must perform while flying an airplane is termed "workload." No single standard means of measuring workload by the aircraft industry exists today, and it has only been since the early 1960's that detailed workload studies were even attempted for new aircraft designs. The lack of a single standard procedure for workload demonstration is not necessarily disadvantageous, and it is certainly not the same as there being no useful methods. As a result of analysis by the applicant for certification and the FAA, a standard method is developed for workload evaluation of any given airplane. Of necessity, this agreed method of assessment is keyed to the unique features of the cockpit in question -- those aspects that have been changed from configuration in baseline aircraft must be most exhaustively studied. Therefore, it was necessary in the case of the B-737, later models of the DC-9, and the subsequently developed wide-body aircraft (B-747, DC-10, and L-1011) to develop methods of measuring crew workload and also to determine what the acceptable limits for the pilot's workload might be. The methods specified

for the B-737 were selected as appropriate to the particular changes incorporated in that two-person flight deck design. Subsequently, those methods were refined, changed, and grouped in a different overall procedure for the B-747, a very different type of aircraft with a far more complex cockpit design and a crew of three, and which could benefit from B-737 experience, military programs experience and workload research programs.

While the 1965 recodification of FAR 25.1523 emphasized that workload was the standard to measure size of flight crew, aircraft manufacturers and the military services had studied means of controlling workload for many years. For example, Harvard University's Professor Ross McFarland, a world leader in human factors research, was engaged by Boeing to fly in Pan American Airways flying boats before World War II to study pilot workload (see reference 10). McFarland's investigations had nothing to do with airworthiness certification, but were aimed at safety and operability questions raised by the manufacturer and operator. Continuing on down from the flying boat days to the period of rapid post World War II airline development, further studies were made of cockpit equipment simplifications and improved instruments for low visibility operations. The addition of Appendix D did not materially change the requirement for determination of workload acceptability. It did, however, elaborate the types of things considered when determining complaince with minimum crew requirements.

A pilot flying an aircraft becomes a part of a man-machine system. Certain specific functions that are to be performed in a man-machine system are most appropriately done by man (manual system), others require both man and machine interfaces (semi-automatic systems), and still others are most appropriately handled by the machine (automatic system). There is a need to make a decision on the type of system to be used early in aircraft design since the resulting design effort will depend on the decisions.

It is impossible for a pilot to describe accurately the workload required to fly an airplane, just as it would be for an average person to describe the workload associated with driving a car. This is because there are many and varied conditions and factors that prevent the necessary functions from being accomplished in the same order or exactly the same way on every occasion. In addition, there are several variables such as mental workload, crew attitude, personalities, environmental conditions, familiarity, fatigue, etc., which make an absolute and complete prediction of workload under all conditions impractical if not impossible. The scope of such variables is sufficient to make it difficult to identify and resolve key problem areas when pilots rate workloads as high, or to clearly identify the cause when workload becomes "acceptable".

However, certain properties that are known to be characteristic of workload can be measured. Head movements, eye movements, task-time, and time motion are examples of these. By the use of these properties in comparison type study between two airplanes, system designs, or arrangements, a series of quantitative measurements of aspects of workload can be obtained.

The first step in design of a new flight deck and preparation for crew workload evaluation is the formation of a design team. Initial design objectives will cover planned crew size and incorporation of features that

have been found in earlier designed aircraft to provide comfort and convenience to enable efficient crew functioning for the designated number of crewmembers. In many cases, the initial plan will also call for increased automation in avionics, flight control systems, and overall flight management systems to take advantage of the increased capability of computerized devices.

It is the responsibility of the design team to develop a plan for flight deck design and workload documentation. Since the overall decision verifying the acceptability of flight crew workload must apply to the whole range of anticipated flight operations and both normal and contingency conditions, it is axiomatic that no one individual test method, prior to flight test, can define workload absolutely. But in combination and applied at appropriate stages of cockpit development, the various part-task and simulation methods can serve to answer key questions adequately to avoid later retrofit or change in crew complement particularly as to the workload impact of identified changes from baseline designs.

The planned crew size will obviously affect the degree of detail and type of workload testing to be accomplished. Should the new aircraft development constitute a derivative of an existing one, or one with no change in basic crew functions from earlier, in-service aircraft, the workload program may be relatively less detailed. If, in contrast, a reduction in crew size from that employed in all the applicant's previously certificated aircraft is proposed for a new design, the workload evaluation plan will have to be sufficient to prove that the new aircraft is safe.

As a general principle, it is recognized that an aircraft is determined to be safe or not safe. There is not a continuous scale of safety paralleled by degrees of workload. Hence, the workload documentation plan must be adequate to provide for those data which are necessary to show that no unacceptable peaks in either pilot or other crewmember workload will occur. In any case, the applicant is responsible for producing a workload documentation plan and obtaining FAA approval of the plan regarding its completeness and objectivity.

2.3 Crew Workload Studies Made During Aircraft Design

Aircraft design is, of course, a continuous process. Under inhouse support, manufacturers' design teams are continually examining potential variations on existing aircraft types, applications of new technology that may make development of a new aircraft feasible, and customers' evolving aircraft requirements. Furthermore, present production aircraft are evolving with improvements and occasional changes of cockpit equipment. In exceptional cases, an opportunity exists to develop and flight test a radical departure from prevailing cockpit patterns. An example is the National Aeronautics and Space Administrtion (NASA) Terminal Configured Vehicle (TCV) program, under which Boeing modified a B-737 to provide a second cockpit with electronic flight instruments and other novel systems (see references 11, 12 and 13). Other examples have been noted in military developments and experimental flight programs conducted by transport aircraft manufacturers. The result of all of the design work summarized above is the accumulation of a growing body of actual flight test information available to the initial design team starting work on a new transport

cockpit. The starting place, then, for crew workload inquiry in the case of a new aircraft design is a body of data from preceding research including results of the most credible type, such as results of actual flight (Figure 1).

The special workload studies usually made during the design phase of a new aircraft program are generally of four types: operating scenario and procedure comparison, task/timeline analysis, task/motion computer studies, and functional mockups. Each of these types is normally based on a comparison between an initial concept for the new flight deck and the existing design of a proven in-service aircraft flight deck.

Operating scenarios and procedures from the pre-flight through the flight phase profiles and into the post-flight operation are assembled for the new and comparison airplanes. It is then possible to analyze the complexity and number of procedures for each phase of each airplane operation.

Meaningful task/timeline analysis becomes practical as the preliminary design is coupled with the operating scenario and procedures. Task listings, then timelines are developed to follow the scenario, based on the design concept. Timing, complexity and compatibility of tasks are appraised by the analyst, and, as required, the design concept is refined to avert conflicting demands for crew performance. Both manual and computer methods are used to compare required task performance time versus time available for performance. When the comparison indicates that required performance is only marginally adequate, further refinements are accomplished in those design details which dictate task demands until results indicate reasonable workload levels have been achieved.

Task/motion computer studies are sometimes found to be valuable at the beginning of a program since they can be easily adapted to a computer which then allows many variations to be evaluated in a minimum of time. Before doing task/motion computer simulations, it is obvious that the cockpit or the options for cockpit features, including display and control details, must be well enough defined so that they may be programmed adequately. Hence, if the new design program is not starting with a well defined prototype, such as might have been evolved in previous flight tests, these computer studies will require layout sketches or drawings at a minimum. Unlike the preceding procedures study, the computer study is based on summing actual movements. In capsule, the method may call for the selection of representative flight segments involving high workload. Analysts or trained observers then define and analyze the profile that includes all the actions required by each crewmember over the flight segment. Since early in the design cycle, actual measurements of time required to accomplish the tabulated operations may not be available, reliance is usually placed on measurements from previous reports or a company data base developed from earlier tests. (see reference 14). For confirmation, detailed recordings such as time and motion photographs or video records may be made of actual pilots operating in the selected high workload environments in the comparison aircraft.

With the profile and the required crew tasks established, it is possible by means of a computer program to calculate the amount of head/eye motion

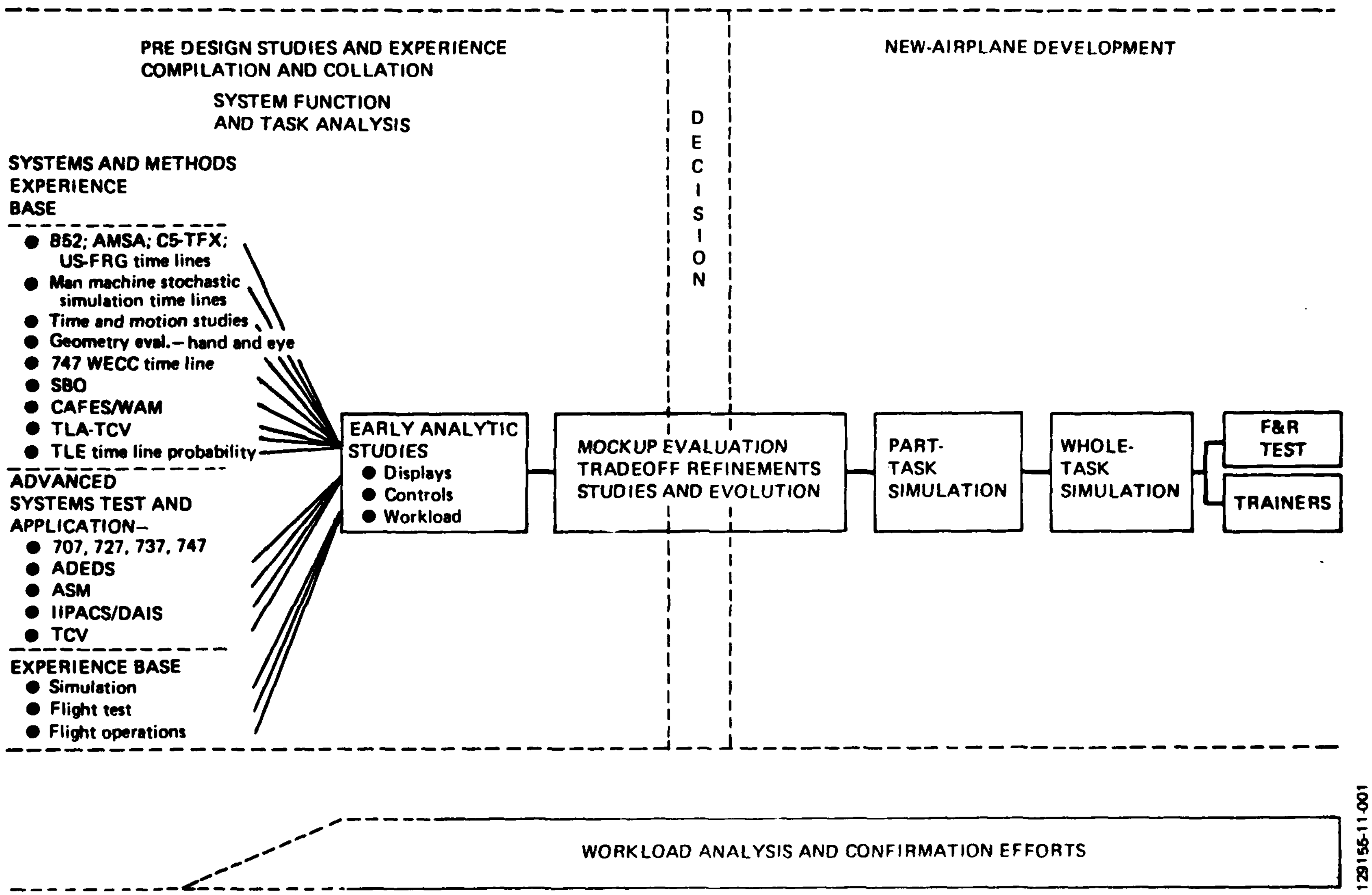


Figure 1 Representative New-Airplane Progression for Workload-Related Appraisals

D. PARKS 2/4/81

in terms of degrees of movement and the amount of hand motion in terms of inches of travel to perform every required operation for each crewmember. Normally, these calculations would not be made for every subsystem but would be concentrated on new features or critical relocations. Using this technique, detailed studies are made to determine the optimum location of some of the critical controls such as fire switches, the autopilot panel, the flight director mode selector, and system panels. Comparison studies are then made for various proven, in-service aircraft under the assumption that the crew workload in those aircraft is reasonable and appropriate. (In the opinions of many authorities, this provides the best of the available controls). To the extent that the comparison aircraft includes features that can be juxtaposed to those of the new aircraft, comparison on the same set of flight scenarios may be used. Unfortunately, when there is a radical departure from earlier technology, such as tests of new systems that far exceed the capabilities of earlier systems, such a head-to-head comparison cannot be made this way.

Mockup studies are the fourth type of design phase investigation of crew workload. While this is not usually done for a three crewmember cockpit derived immediately from a similar design, in the case of a widely changed flight deck concept, an animated or partially operable control cabin mock-up may be constructed and furnished with actual equipment and controls where possible. The instruments are lighted; many are made operable through simulation computers, and many system controls are made functional with switch and light logic duplicating expected system performance. Such a mockup may also be equipped with operable interphones.

To make the mockup an effective development tool, it is evolved into something approaching a procedures simulator or part task simulator by being equipped with an extensive system for duplicating faults in the various systems. The pilot's main attitude indicator can be varied at random from a control console outside. The pilot then provides control wheel movement to correct back to the desired flight attitude. While the pilot is thus occupied with a recurring requirement for control wheel inputs, the operator at the console outside the mockup simulates faults with appropriate indicator lights and aural warnings. This provides an accurate reproduction of single and multiple malfunction conditions that might actually be encountered in the aircraft. The subject pilots in the mockup then take the appropriate action, the evidence of which is monitored from the external console. By using this partial cockpit simulator, it is possible to evaluate certain types of equipment during systems malfunction, singularly, in combination, or during periods when other flight activities are underway.

For a two-crewmember design, typically at this stage, progressions from the simple to the animated cockpit may be used to demonstrate to the FAA a complete flight, including communication and navigation functions. The mockup is also used in presentations to customer pilots and pilot groups. Many valuable comments and suggestions are made by the pilots who evaluate the animated control cabin, and a large number of these suggestions may be incorporated into the basic design.

If the new aircraft cockpit design is conventional, it is known what properties are required to make it acceptable to the crew. In that case, it is

not necessary to get new data. Such last fragments of required changes or adjustments that may be identified as necessary can be made during the flight test phase.

2.4 Crew Workload Studies Made After Aircraft Design

In addition to possible continuation of the three types of studies developed in the design stage, additional studies are added between design and the first flight test. Usually, the functional mockup is updated to reflect the latest design, and preliminary procedure and checklists are finalized.

Each system is evaluated as to failure modes, indication of failure, recommended response to failure indication, and consequence of failure to respond to a warning. This is to assure that the crew need never provide constant attention to a system and that system controls and indicators are given a priority of location consistent with the importance of the particular system operation under normal and abnormal condition.

Multiple malfunctions that could compound abnormal and emergency functions, such as the loss of a single engine and the opposite engine generator on a single flight, are evaluated. For example, one Boeing program evaluated the effect on crew workload of all dual malfunctions with a probable occurrence equal to or greater than one over the estimated fleet life of the aircraft.

Using the advanced mockup and standard procedures, the accessibility and conspicuity of all necessary instruments and controls are evaluated. Pilots ranging in height sufficiently to represent the prospective pilot population are used in the study. Evaluations are made under bright sunlight (simulated with high intensity lamps), normal lighting, and medium to full darkness. Studies that are usually conducted include peripheral vision of instruments, visibility and readability of instruments, and capability of each pilot to see and reach required controls.

As a result of these studies, additional refinements may again be made to the design. Addition of anti-reflective coating on the glass and improvements to visibility and readability of many instruments might be made. Dimming of all indicator lights in the pilots forward and normal peripheral field might prove desirable, and features of the master caution and warning system might need revising.

Two techniques, task/time simulations or probablility distributions, can now be further developed.

For many years it has been customary to attempt to express workload numerically in terms of a time line to show the value or extent of the workload (expressed as a percentage) at any instant in a given mission. It should be kept in mind, however, that as skill increases in working in a complex situation, the operator or pilot does not devote all of his attention to any one action. The pilot can, in fact, look outside and receive information at the same time that he is making control movements or receiving aural information.

The output of the computer replication of normal and contingency conditions over a large sample of flights is usually compared to a similar computer summary for a proven, in-service aircraft. The degree of success in attaining the workload design objectives may be evaluated by examining these two computed task/time probability distributions to demonstrate, for example, that the likelihood of both pilots being fully occupied at the same time is minimal.

The purpose of either of these two techniques is to demonstrate:

1. At no time does the workload on an individual crewmember exceed the maximum of what he is known to be capable of handling in the comparison aircraft.

2. The captain's workload only exceeds an accepted percent (e.g. 50%) for brief periods such as during the takeoff roll when full concentration is required of all aircraft.

3. The nonflying pilot has workload exceeding 50% for only brief periods of time.

4. Workloads are comparable to those in comparison aircraft and show no great peaks that are avoided in alternate design.

The early development of a flight simulator representing the cockpit/systems of new aircraft developments also helps in workload evaluations. Here, as in the mockup study, malfunctions can be introduced by an instructor's control panel. The flight simulator referred to here is an engineering simulator that is not exactly representative of a final design, usually is built to permit easy modification, and is not usually equipped with either a motion-base or an outside visual attachment. The ultimate, "training level" simulator obviously cannot be built until the aircraft is fully developed, and at that point, the test emphasis shifts to actual flight test. (See section 3.4 for a more complete discussion of the engineering simulator.)

External vision studies are made to determine the amount of outside visual field available for the pilots compared to proven, in-service aircraft. Also, if it is related to a critical design change, the potential contribution of an observer or third crewmember may be assessed. The relationship between available external vision and collision avoidance can then be computed for different speeds and for the separate crewmembers, again, if crew size and collision prevention by visual detection is considered an important issue. The available visibility that a crew member has is a function of the shape and size of the window and the eye distance to the window. In engineering terms this is measured in steradians or the solid angle subtended. From the external vision measurements, charts can be developed to document that the new flight deck design meets reasonable criteria of external visibility.

Manufacturers of FAR Part 25 approved aircraft have developed, and used successfully, various additional computer modeling studies. Geometric

data are used to evaluate the visibility and accessibility of items in the cockpit prior to final construction of the aircraft. Angular movements and changes in linear distances for the eyes and hands as they move to perform aircraft procedural tasks during flight are calculated. Combining these data with programs covering task sequences allows determination of angles from the eye to points of interest and fields of view during crewmember procedural activity. Changes in angles and linear distance can be summarized for mission segments, showing how much eye fixation must be shifted to operate the aircraft. Eye/hand coordinated motions and procedure execution times are included in various programs of this sort so that overall eye movement workload results from arrangement and procedures changes may be evaluated quantitatively. Accessibility-reachability is also tested in the three-dimensional mockup. Available analytic tools such as computer modeling methods are also now available.

In summary, the final culmination of the various workload studies made after the basic aircraft and flight deck designs have been fixed is a comparison of crewmember workload with that of reference aircraft. While authorities agree that total workload cannot be measured in such a way that a numerical safety factor can be derived, the various workload measures obtained in simulation, computer, and analytic studies can be compared between a new flight deck design and proven, in-service designs. No single comparison, as a simulator study alone or a task/time probability distribution alone, is considered to provide an adequate comparative evaluation. Taken together, however, the results of the several types of studies have proven to be predictive of the real-world workload effects.

Ultimately, knowledge of the range for optimum level of workload may become desirable. This optimum would be the kind and amount of workload that keeps the pilot in touch with all critical aspects of the ongoing activity and keeps him ready and able to take over and fly manually. There can be too little as well as too much workload. The goal of cockpit improvement then, is not simply workload reduction. When a pilot flies manually, he must fly accurately and efficiently and safely. A person cannot instantly go from a very low workload to fully integrated complex performance unless the situation has kept him in the loop, at least mentally. It is a reasonable goal to seek to reduce subsystem workload, but that does not mean that the goal should be to further reduce cognitive workload to a similar low level.

3.0 Procedures - During Aircraft Design

3.1 Flight Crew Task Analysis

Since the flight decks of today have evolved as a result of experience gained over decades, there is a great deal of continuity in basic design and required flight procedures. Some minor portion of this regularity of configuration and operating principle is actually required by the FAR's, but much more of the detailed similarity results from the universal desire in the aircraft industry to preserve, as far as possible, conventions and practices which have been long established. While available technology might permit nearly total change, and such revolutionary designs are sometimes encountered in special-purpose military aircraft, it is agreed, in the airline context, that change should be made only for good cause. There should be a planned positive advantage in making changes, and the degree of actual attainment of the planned gains should be openly assessed during aircraft design stages, in contrast with, for example, losses in training transfer.

The case for change is usually clearly stated in the initial aircraft program plan. While keeping the evolutionary character of useful changes in the forefront, the company design team will seek to advance the state-of-the-art in flight deck design to attain, for example, more balanced crew workload, improved flight safety through improved visibility, simplification and increased flexibility, and greater reliability of equipment components and systems. Pressure for change comes from recognition of the potential gains in areas such as these and also from the increasing difficulty of accommodating the instruments and controls necessary for operation in a future environment. New navigational facilities, such as 4-dimensional area navigation, new flight management systems benefiting fuel-efficient flight planning, and new safety systems such as ground proximity warning or added collision avoidance devices, must be fitted into the overall design. As an example, there was no good place for the large area navigation map display when initial system implementation began. The result in most airline cockpits was placement to the left side of, or behind, the captain making ready reference awkward. In a new design, new display space must be located in the convenient, forward area.

Congruent with the discussion in previous sections 2.2 and 2.3, it is important to recognize that the aircraft design process may go on for a considerable period of time and result in the accumulation of substantial data and the making of important cockpit configuration decisions before the formal announcement of a particular aircraft program or even the engineering "Authorization to Proceed" (ATP). ATP is a milestone established by some transport manufacturers to mark the point at which the company has completed sufficient trade studies to define a significant potential market and has a concept for the new vehicle that justifies the large expenditures necessary for final and detailed engineering design.

Both Boeing and Douglas assert that advanced engineering studies of propulsion systems, aerodynamic developments, structures, subsystem advancements and the like constitute a never-ending process bridging one aircraft

program and its successors. During this phase, issues are defined in discussions with potential customers (trade studies) and various concepts for new aircraft are discussed and estimated as to costs, weight, and prospective performance. There is, of course, a crossfeed of information from military and civil aircraft design teams. Many such preliminary military and NASA design efforts do not result in production programs, but continuing advanced engineering efforts in the cockpit design areas have led to early workload concepts and to many currently used techniques, and continue to provide refinements in workload assessment procedures.

Workload evaluations conducted during these advanced development stages are not necessarily part of the workload documentation plan prepared to show compliance with FAR 25.1523. As illustrated in the sample cockpit development plan shown in Figure 2, extensive analytic efforts (including preliminary workload analysis) as well as laboratory tests, mockup evaluations, part-task simulations, and even full-mission simulations will be conducted prior to completion of the plan for a cockpit production configuration. As shown in the sample plan, even aircraft ground tests and aircraft flight test are included in aircraft design configuration.

A substantial body of flight test data may precede aircraft design. Early flight tests must be conducted in modified existing aircraft, since the new aircraft has not been designed, much less manufactured. In an aircraft design following this paradigm, a wide range of workload evaluation procedures will have been employed before any attempt is made to establish a FAR 25.1523 compliance plan or to negotiate acceptance by the FAA of that plan. It is clear, then, that there may be two stages of workload study: evaluations conducted during aircraft design and tests performed on the one selected production configuration to show that the concept is safe when operated by the specified crew complement.

Size of flight crew will be one of the subjects covered in trade studies and will be stated as one of the starting assumptions in the initial design plan. Typically, the years elapsing since the manufacturer's last new aircraft design will have been used to develop additional design objectives in considerable detail and to build up a data base on crew duties in existing, potential comparison aircraft. In the case of the DC-9-10, the DC-8 nose shell provided a starting point for the design of the cockpit as both aircraft share a common cockpit enclosure.

The development of the DC-9 cockpit reflected the improvements derived from Douglas development programs, from inputs provided by airline flight and engineering personnel, and from inputs of an Air Line Pilots' Association (ALPA) committee. The main Douglas effort was directed by a cockpit committee composed of pilots and design engineers, with pilot responsibility at the project level. This approach ensured that crew requirements were constantly considered as a part of the basic design requirement.

Specifically, crew checklist procedures were written prior to system design so that the system design conformed to the crew procedures rather than the procedures being written around an already designed system. This process also considered emergency procedures, and the systems were designed to accommodate those emergency procedures.

This design effort was supported by system research tests on display systems, time and motion studies of DC-9 procedures, and an analysis of the crew tasks anticipated in DC-9 operations. The system research tests were mainly applied to developing the centralized master caution and warning system. A DC-9 functional mockup was developed into a procedural trainer for use as a working tool to simulate actual conditions. This mockup made it possible to test and research procedures relative to the aircraft systems design. To analyze those procedures relative to crew function, two measurement techniques were used to obtain data. The first was the chronocyclographic technique, a time exposed still photograph of subjects performing a task with small flashing light bulbs attached to their index fingers. The second was a micromotion technique, an accurately timed motion picture.

A full and complete profile of the tasks and task sequences required to operate the aircraft in high workload flight segments, such as takeoff and approach, is developed from the manufacturer's previous detailed pilot task analyses, supplemented by the information reflecting changes in planned design. This profile provides a "job description" with detailed itemization of each required step in the new aircraft's standard operating procedures (SOP's). Everything that must be done, by whom, when, how often, and in sequence and combination with every other necessary action is enumerated.

The second part of the data base that is necessary at this stage of design is the event time for each identified action in the flight scenario action profile. Dating back to a 1962 study by the American Institute for Research and other classic experimental analyses of the required task times for operation of specialized equipment, manufacturers have developed substantial data bases of this kind as well (see reference 14).

The result of this initial setting of flight deck program objectives, initial design planning, checklist and procedures development, and updating of the task time data base is the accomplishment of a flight deck system description with accompanying flight crew task analysis. Since the system design interacts with procedures changes, these activities are conducted in parallel and with iterations. For example, the introduction in recent aircraft of advanced flight control and management computers, trending toward the so-called "push-button" airplane, means that system design and procedures analysis will track together, rather than in sequence. With this phase of the design cycle completed, actual comparison of crewmember workload with that in a reference flight deck, and measurement of workload in one design version versus another, can proceed. In this connection, it should be noted that in some programs there have been multiple cockpit mockups, and in the case of the DC-10, there were two simulators in use at the same time, one cockpit tied to the "iron bird" and another employed in tests of flight management system elements. A revised method of handling altitude selection, area navigation procedures, the revised caution and warning system, and similar subjects were decided based on the simulation tests of alternative proposals. Whether or not several simulators or mockups are used, it should be kept in mind that the goal is to establish a single cockpit design prototype and not to create alternate designs for

final run-off comparison tests. After attaining the most suitable single prototype design, the comparisons are made against a cockpit that has been previously proven in actual line service. As stated above in Section 2.3, there is a significant limitation in this comparison in that radical departures and improvements in automation in the new design may not be paralleled by features in the baseline design, thereby making direct comparison unfeasible.

3.2 Operating Procedure Comparison

To set the stage for an initial full-task comparison with a reference flight deck or alternate design concept, the operating procedures from preflight, flight phases, and post-flight operation are assembled for the new, and for the reference design, and all checklists and individual system procedures are paralleled so far as equivalence of features allows this to be done. It is then possible to analyze the complexity and number of procedures for each phase of each airplane's operation.

Many valuable inputs regarding operating procedures were made by pilots employed by the manufacturer, pilots of customer airlines, FAA pilots, and by pilot association evaluation committees to the design of various new cockpits such as that of the DC-9. In addition, proposals were made by consultant pilots to state requirements that would aid in reducinq crew workload going beyond the control of the manufacturer. Some examples of these requirements have been that there should be absolutely no required inflight paper work in the cockpit; another was that there were to be no company radio contacts except in an emergency or at the pilots' discretion. It may be noted that this issue continues to be a problem area into the 1980 decade.

A principal design objective for the B-737 was to eliminate inflight crew actions necessary for the main systems (fuel, air conditioning and pressurization, electrical, and hydraulics) since the monitoring and operations of these systems is the primary job of a third crewman in a certificated 3-crewman aircraft. It will be recalled that the major goal in improved designs for workload control, as summarized in Section 2.4, is to reduce subsystem procedures and required activities. While no one knows exactly what is the optimum total workload, there is general agreement that keeping the flight crew actively engaged in flying the aircraft is greatly preferable to permitting systems such as those monitored and adjusted by a flight engineer to command major attention. The result of this simplification development is shown in Figure 3 for the fuel system as a representative example:

The 737 fuel system is a simple two-tank (shown) or three-tank system which requires no inflight fuel management and no fuel dump system since the airplane structural provisions provide for landing at a near takeoff weight and accommodate 1500 pounds inflight fuel imbalance.

The normal 737 fuel procedure is to turn all fuel pumps on prior to taxi and turn them off after landing. If the 737 is provided with a center

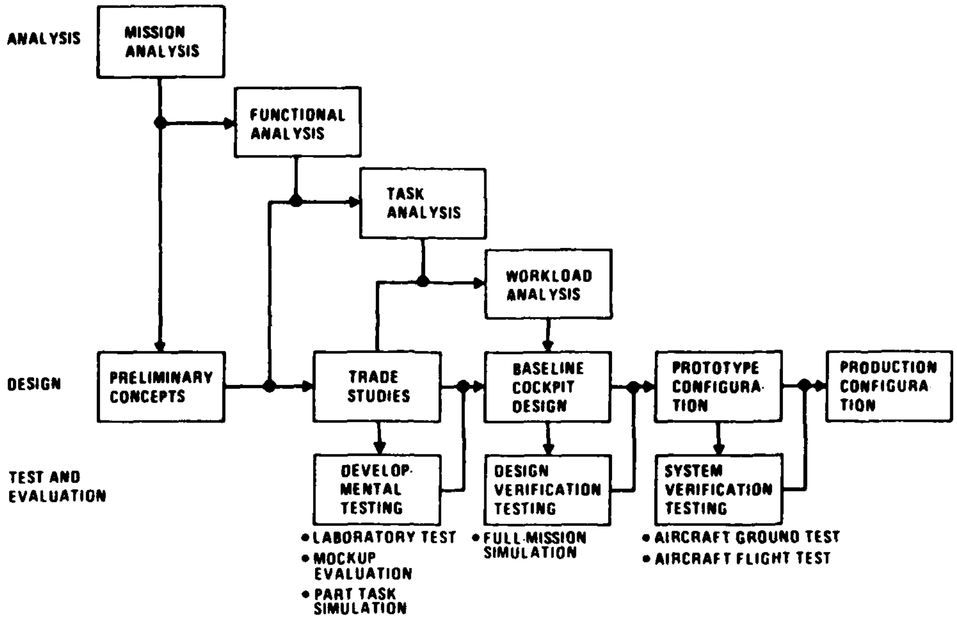


Figure 2 Sample Cockpit Devlopment Plan.

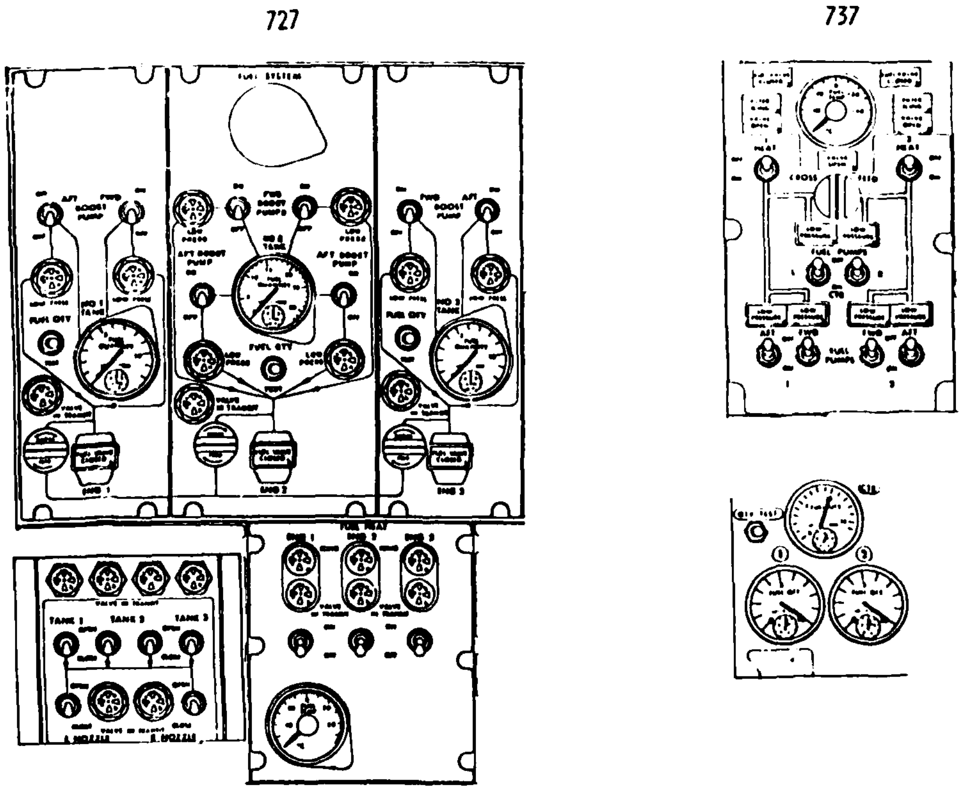


Figure 3 Fuel System Comparison (B-727 vs. B-737)

tank, a master caution indication for fuel will show when the tank is empty. The two switches for the center tank fuel pumps may then be shut off. When fuel heat is required during flight, the filter icing/master caution light indicates to the crew to turn on the fuel heaters which are then automatically timed and shut off.

The 727 system management requires tank to engine operation during engine start, takeoff, and landing. After takeoff, all engines are operated from number two tank until all tanks are equal, and then they are returned to tank-to-engine configuration. The quantity of fuel in Tank Number 2 must not be less than the quantity in Tanks Number 1 and 3. Quantity indicators must be monitored to prevent landing with overlimit imbalance or overweight conditions. The three crossfeed valves and eight boost pump switches are used in normal inflight fuel management, and the fuel dump system is used to prevent an overweight landing. The fuel heat systems are operated one minute every 30 minutes when the icing light indicates a requirement. There is no master caution system, so the 26 lights and 5 fuel gauges have to be monitored periodically during flight by the flight engineer.

Early in the design stage, it is characteristic to draw up a table of crewmember duty allocations for the new, and for the comparison flight deck and to estimate the impacts on crew workload and operating safety of each difference, often using the computerized methods. This process makes explicit the expected results of each change in equipment, location, or operating principle and thereby guides the selection of experimental comparisons needed to verify the attainment of design goals.

3.3 Task/Motion Computer Studies

As a characteristic example of task/motion computer study methodology, the Douglas Aircraft Company developed a computer study method for the DC-10 that provides the capability to measure flight crew workload as it is affected by alternative crew station layouts, controls, and displays (see reference 15). This method is very much the same as the one developed earlier by Boeing for the B-747 (and which has since progressed through several refinements; see references 23, 24, and 25). As stated in Section 3.1, whether dealing with the initial or baseline cockpit design or an evolution toward the prototype configuration, these studies are made on a single design, not on all possible variations. Within the areas of uncertainty of that single design, perhaps arising from pilot observations in mockup study of visual interference between the control yoke and displays or relating to cross-cockpit visibility questions, several possible relocations, size changes, or geometric adjustments may be considered as candidate variations on the single design.

The computerized technique concentrates on design factors under the control of crew station designers and provides for quick and low cost iteration of alternatives. The program computes workload as related to specific equipments and systems, permitting special attention to be given to high workload items during the early development of concepts and hardware before simulation is available. The technique and program are also applicable to integrated displays including those where programming to meet information requirements is an element. Since it actually takes from two and

one-half to three years to detail a new flight deck design, specifying what is there, how big, and all the other details necessary before possible construction of a prototype model, there is no attempt to design two or more major variants. A single improved, potentially acceptable design is the goal.

The analysis is based on a typical flight mission scenario constructed to explore the expected operational envelope and to exercise the aircraft displays, controls, and systems in a sequence and time frame typical of the more demanding operations planned. The primary measure is the ratio of the required performance time to the time available within the time constraints derived from a specific flight scenario, supplemented by hand movement action and distance data. The operating procedures are detailed for computer handling in a way that relates a single workload element to a single piece of equipment with the equipment coded by its location. The times for completing specific acts in the cockpit are developed by detailed analysis of each task and its associated equipment using standard action and reading times. The hand movement and distance data may be evaluated using a full-size design aid.

The basic definition of workload employed in this concept is as follows:

> Flight crew workload is defined as the ratio of the time required for crew-equipment performance to the time available within the constraints regulated by a given flight or mission. For comparison purposes, workload is expressed as an index having a numerical scale of 0 to 100. The value of "100" represents the circumstance of the crewmember utilizing the total time available to execute the tasks required for safe aircraft operation over a given route. Conversely, the value of "0" indicates that there are no crew actions required.

This definition can be expressed as follows:

$$\text{FCWI} = \frac{T_R}{T_A} \times 100$$

where:

FCWI = Flight Crew Workload Index
T_R = Time Required
T_A = Time Available

Thus, time is used as the basis for the evaluation of flight crew workload in a manner that is consistent with time and motion study concepts employed in traditional industrial engineering practice. Each task input to the

model describes the interactions between the flight crew and operation of various controls and displays that must be actuated or observed for safe control of the aircraft during the flight scenario.

To cover abnormal conditions and system malfunctions, which often are the peak workload causal factors, the flight scenarios are rewritten to include realistic events. In this way iterations of the computer calculations provide information on the relative workload impacts of various possible eventualities.

Computer assisted processing is employed to categorize and analyze data. Several types of analyses are then performed to summarize design comparisons for various individual and combined aspects of crew workload, including an external vision availability analysis, an equipment interface workload analysis, and an overall comparison with the reference aircraft.

Analysis of time available for scanning the outside environment is important and is determined by the time required for internal cockpit tasks such as viewing displays and controls during the course of the flight. It is not possible, of course, to separate command decision time, which is variable and unobservable, from external viewing time. And, as noted in Section 2.4, the skilled pilot looks and thinks at the same time. Hence, the perspective of a comparison from one design to a baseline cockpit must be maintained and no undue credence placed in the computer generated times as absolute or "true" real-world representations. The computer program examines data in the task file, sorts the data, and prints out the external vision time available for both crew members as a function of milestone start times and duration. In addition, for the Douglas approach, a routine is provided for combining the Captain's and First Officer's external viewing time in a graphical format so that total external vision available to both crewmen may be ascertained throughout the flight.

Body channel workload represents quantification of the overt physical actions taken by the flight crew to operate the aircraft. Each crew member is linked with his work environment primarily through physical, visual, vocal, and aural channels with respect to sensory/motor input and action. Involvement in work tasks may, therefore, be measured in terms of the level of induced activity of these channels.

These programs thus determine the detailed work allocation on a task-time basis for each crew member considered as a five-channel input/output subsystem. It reflects a composite of the physical actions, reactions, and perceptions necessary to fly an aircraft along a prescribed flight path. The flight crew workload analysis thus produces results in tabular and graphic format reflecting the combined duty cycle of total visual, aural, vocal, and body extremity operational activity.

The overall cockpit activities necessary to achieve flight path milestones are then summarized by individual body channel as a function of the percent of time required to time available, and a plot is made of peak values which might be indicative of potential overload.

The concept of the crew workload in interfacing with equipment represents the total percentage of available time that is utilized by the flight crew to control the aircraft during flight. The computer program sums each of the individual crew task times and relates this to the time available within each of the events associated with a particular flight. Since the program treats all body actions as occurring serially and does not reflect human capability for simultaneous use of two or more body channels, the workload values computed for an individual aircraft can, therefore, be considered conservative. In addition, these measures of workload are combined on a time-weighted basis to provide for an assessment of workload for each flight segment as well as an overall average for the entire flight, i.e.:

$$\text{Segment Time Weighted Workload} = \frac{(WL_{XA})(T_{XA})+(WL_{XB})(T_{XB})+\cdots(WL_{XN})(T_{XN})}{T_{XA} + T_{XB} + \cdots T_{XN}}$$

WL_{XA} = Event XA Workload (T_R/T_A)
T_{XA} = Time Available for Event

A weighted formula is used to combine the several different measures into a single number. Such a single number has only an arithmetic basis, since there is no good rationale for the combining process assuming, as it does, the single channel hypothesis of pilot activity. An overall time-weighted workload for the full flight segment is computed in a similar manner.

In the case of a typical application of these design stage task/motion computer studies, the major purpose is to provide an objective basis for choice among design alternatives within the single overall configuration, and to verify that preliminary design plans do appear to accomplish their stated objectives. As a result of such studies, a larger number of alternative layouts and equipment choices would be expected to be narrowed down to a selected few that could then be tested in flight deck mockups in a further iteration of the process described above as originally developing some design choices from pilot analysis of an initial mockup.

3.4 Functional Mockups/Simulators

The accessibility, ease, simplicity, and conspicuity of controls and malfunction indications and the extent that such devices direct the proper corrective action are considered major workload factors and, as such, are listed in appendix D of FAR 25 "Criteria for Determining Minimum Flight Crew." (See reference 2). One of the basic design criteria of a new and improved flight deck may be, for example, to group system controls together so that corrective action required during a malfunction, perhaps a switch or knob actuation, occurs in a logical place adjacent to the indicator light calling out the malfunction or, otherwise, in a position convenient to view any necessary gauge or other indicator of restored status.

Comparisons between alternative locations, and groupings of such system controls, can be made by examination of diagrams, but for final design choices, it is considered best to make such tests in a functional mockup, a lighting mockup, or a configuration control cab.

It is ordinarily the case that the initial design concepts will have been elaborated in detail and examined analytically by pilots and reviewed in overall flight deck layout plans. At this stage, most design features will have become well fixed by consensus agreement among the outcomes of these procedures, but there will probably remain some areas of uncertainty, particularly when space is at a premium and there are competing claims for the "goldplated" locations directly in front of pilots and at the panel eyebrow. Furthermore, the overall configuration and relationships among overhead, pedestal, and front panel locations can best be visualized and tested in a three-dimensional environment.

In many new aircraft design programs, the initial mockup is a simplified version with paste-on pictures of most displays and many operating components substituted from an earlier design. As more and more design decisions are made, the realism and detail of the mockup is upgraded so that final design reviews are made in a pilot's environment that looks, sounds, and feels very much like an actual aircraft, although the flight equations are not yet installed in a simulation computer to permit practice of actual flight maneuvers.

One set of questions best studied in a mockup was noted in Section 3.3, that having to do with visual yoke-interference, cross-cockpit visibility, etc., but there is another topic that must be treated in physical tests. This is anthropometry, including factors of pilot size, strength, control reach, and adjustment range. Since the present trend is toward more diverse sources for pilot populations, particularly including females who are, on an average, smaller and less strong in arm and hand power than men, it is particularly important to test the suitability of dimensions and required forces.

Manufacturers frequently use "soft" mockups quite early for engineering trade-offs in a comparative evaluation of two or more alternative crew station designs in side-by-side mockups. Given a significantly important question to answer, this might continue to the later phases of design development in dynamic cockpit simulators. Mockup comparisons of this kind have used pilot questionnaires, and with the dynamic flight deck cabs, primary and secondary task scores have been employed.

During development of the early turbojet aircraft, such as the B-707 and the DC-8, and even the B-727, the primary methods of workload evaluation in design were relatively simple. Mockup reviews by experienced pilot groups and pilot-engineers were used extensively. This certainly does not mean that these programs were not done well. It is a fact that, despite the development of computer programs for detailed computation of comparative workloads and other observational techniques such as video recording or time and motion photography, primary reliance is still placed on review

by experimental test pilot groups. The validity of laboratory and synthetic methods utilizing computers and standard data bases is considered to be greatest for specific head-to-head comparison of a design feature to a standard or baseline design. For an indication of actual acceptability to operational pilots, pilot opinion based on systematic examination of a mockup and detailed knowledge of actual airline operations in existing aircraft is generally thought to be more indicative of what will subsequently be experienced in actual line service. One major difference between earlier programs and the more recent ones, particularly in workload evaluation for a two-crew design, has been the level of documentation required. Following completion of the design, with most design choices made on the basis of cockpit mockup evaluation, the acceptability and suitability of workload was confirmed by inspection and evaluation during initial certification testing and the functional and reliability (F&R) phases of flight test.

During design of the B-737, a more thorough and documented examination was concluded very early to be necessary because of the need to substantiate safety of operation with the reduction in pilot crew from the previous Boeing turbojet crew-complement. Accordingly, a very extensive program was undertaken to evaluate and design the layout arrangement and configurations of the various cockpit controls and displays and of automated features used to reduce workload demands.

Before the B-737, there had been, of course, many transport aircraft with a two-man crew, including most of the common piston engine aircraft other than the larger four-engine types. But, since the B-737 was the first new turbojet transport to apply for certification with a crew of two under the relatively new Part 25 of the FAR's, the required documentation of workload acceptability was more extensive. Now that domestic and worldwide airline experience over the years has validated the conclusion that a crew of two is appropriate for that aircraft, it has been suggested that a future design might properly be evaluated by simply comparing it to the B-737, rather than by conducting a full proving process and comparison to an earlier three-crew design. Despite the logical appeal of that proposal, it may be that the practical reality of acceptance by non-specialists, in the face of criticism that might come from pilot unions, will require future repetition of extensive documentation that two pilots are the correct crew of any new aircraft designed for that crew.

For the B-737 program, new computer modeling techniques were developed to evaluate geometry, arrangement, and task sequencing in terms of travel by the eye and hands; modeling was also applied to appraise functional organization or "efficiency" of the various panels to be operated during normal and off-normal flight operations. A computer model was developed for these studies of normal and contingency procedures using various flight scenarios, associated cockpit procedures, the geometric arrangement of control devices and/or displays, and the associated sequence of task operations to produce summary data on the amount of activity (eyes, hands, etc.) involved in satisfying demands.

In this manner, each procedure was examined and directly compared with similar procedural demands in earlier, existing, and acceptable aircraft,

and improvements in arrangement or in automation were incorporated as determined to be desirable. Likewise, in this time period, the extent of vision requirements became a major consideration in terms of workload demand. A computer program subroutine used the same dimensions, points, and procedures in order to examine visual tasks, the extent of the visual tasks, the visual field, and the amount of eye travel inside the cockpit. It should be noted that each such test and correlated evaluation in the mockup was pointed at a single item of interest and the simulation was simple in nature.

To further reduce the risk associated with certifying the B-737 for a two-man crew, a part-task simulator activity was established almost from program go-ahead. This simulator was used to perform evaluation of crew operational procedures and workload related to the aircraft systems. Throughout the evolution of the final design configuration, evaluations were accomplished for airline and FAA representatives.

Initial construction of the flight deck mockup was based on the initial design with minimum changes from the antecedent aircraft as required to accommodate the smaller, two-man crew. As data were generated by the several computer programs, changes and rearrangements were deemed desirable, and these were incorporated in the mockup which, itself, was becoming more of a dynamic simulator as operating functions were fixed and added. Hence, the mockup/dynamic simulation phase of flight deck design should be viewed as a process interacting with concurrent design processes based on computer models and engineering studies. Thus, a mockup is one of the design starting points, and its evolutionary descendant, a nearly full capability dynamic simulator is one of the last test environments.

At this point, several citations of functional and nonfunctional mockups, part-task simulators, and full-mission simulations have been made. For purposes of clarity, several definitions of how these test environments differ should be stated.

Five stages of evolution from a first paste-up likeness of a cockpit to a "training level" simulator may be distinguished. The basic mockup (stage a) has an approximate dimensional similarity to the proposed flight deck but has essentially none of the new or functioning components. It may be furnished with seats, pedestal, and basic panel configurations from preceding aircraft with changes in paste-on instrument and system indicator faces as appropriate to the new aircraft design concept. This minimum, stage a, mockup is used as a starting place with realistic components added in the case of items that have been identified as critical when they become available as a result of detail design decisions and supplier actions. Pilots may examine procedural sequences and simulate various flight regimens, but neither control feel nor instrument fidelity is provided. Normally, there is no motion system and no outside visual scene.

Engineering design aids or "soft" mockups are constructed in the early design stages to assess crew station geometry/architecture, reach and access to controls, and preliminary display/control arrangement. These mockups are of inexpensive "foam-core" material and are easily altered to reflect changes in design concepts. Control/display consoles are represented by

reproductions of control panel drawings attached to appropriate surfaces. Figures 4, 5, and 6 illustrate the type of construction of these consoles, and show, in addition, two other uses for this class of mockup: (1) To assess the grouping of airborne crew stations (usually military) with a common viewing task (the large display screen); (2) To provide a preliminary assessment of primary display/control entities (the viewing screens of each console are backed up by random access 35 mm slide projectors for simulating various display formats and symbologies in static fashion). Soft mockups are frequently constructed as knock-down displays to be taken to potential customers or users for comment and for use as a sales tool.

When crew station architecture is better defined, plywood mockups of the type shown in Figures 7, 8, and 9 are constructed, since overall panel shapes and architecture are unlikely to change. These mockups may have painted ferrous metal panel surfaces to which display and control items, represented as two-dimensional facades with magnetic materials embedded in the back sides, may be attached, evaluated, and rearranged until an optimum configuration is achieved.

The final evolutionary development of static mockups is to prepare a complete wood construction crew compartment with all external features of panels, displays, and controls represented as accurately as possible. These mockups may incorporate integral instrument and flood lighting in order to assess and optimize the lighting environment. Figure 10 shows the flight engineer's station lighting mockup for a modern wide-bodied transport aircraft.

The second stage is the development of the functional mockup, stage b, which has prototype equipment installed, is lighted, and permits manipulation of certain indicators, cautions and warnings, etc., from an exterior control console.

The next evolution from the functional mockup is the stage c, or engineering simulator. As with the initial development of stage a, the exterior dimensions may not exactly duplicate those of the future aircraft, but here the seats will be arranged so that the distances to the controls and to the indicator panels are realistic. Major systems and instruments are dynamic, essentially as they are in a full performance simulator, but the dimensional relationships are not meant to be exactly representative of the future aircraft and, of course, not all the equations of motion are initially available. One version of the engineering simulator is sometimes called a "developmental" cab to differentiate it from a specific prototype design and to emphasize the experimental potential provided by flexible alteration provisions. Often the engineering simulators are not dedicated to a single aircraft program but rather are shared. The mockups, in contrast, are employed in both development work and trade studies, and are specific to the particular program.

The stage c, or engineering simulator, may be austere in three basic ways: it may have no motion system to provide the pilot kinesthetic feedback; it may have no visual attachment to permit interleaving of in-cockpit and external visual duties; and, it may not include all subsystems. The particular features that set the engineering simulator apart are the ability to

represent various aircraft types and the capability for inhouse change, reconfiguration, and software changes. Figure 14 illustrates one engineering development simulator.

Specialized development and testing of systems often employ the use of part-task/development simulators. A representative of this type of simulation, as used on the DC-9-80 development, is integrated into the Digital Equipment and Technology Analysis Center (DETAC) in the Douglas Avionics Engineering Department. Figure 11 is an overall view of this laboratory. This mockup accommodates the installation of various controls and displays which interact with, and are driven by, the DETAC Varian minicomputer system and peripheral equipment. DETAC may be programmed to simulate various control laws (aircraft responsiveness), conditions of flight (e.g., turbulence), and flight phases (e.g., takeoff, cruise, landing), and drive the displays accordingly. Primary flight controls (yoke and column or stick, rudder pedals, and throttles are provided so that pilot subjects may "fly" the aircraft using the displays for guidance, and measures of performance (e.g., RMS tracking error) are recorded for use in evaluating the display concepts. Figure 12 shows the cockpit mockup with representative CRT-type displays incorporated. Figure 13 illustrates the cockpit view of the screen for the video projector system which provides a rudimentary out-the-window computer-generated forward view. The DETAC cockpit simulation was employed extensively in developing and evaluating symbology for the DC-9-80 head-up display (HUD), and was also employed under contract to NASA to evaluate the space shuttle HUD symbology.

Boeing uses a similar "D" cab philosophy, which has supported company programs from the SST proposal period. The most well known result from research in this cab is the advanced display/control system now flying in the NASA TCV, at Langley Research center.

The fourth stage, moving from the basic, paste-on and functional mockup to the full-performance simulator, is a prototype cockpit, or stage d simulator. The prototype represents only one aircraft type, that currently under final design and, hence, bears an exact dimensional simularity to the final design. In addition, a large degree of realism and representation of the single design will be present including: window size, shape, and location; panel dimensions and distances; pedestal design; control handles with selected shape and color coding; input keyboards and all special features of the new aircraft design such as flight management computers, advanced navigation systems, and integrated caution and warning systems.

Subsystems and display features that are unchanged, or revised only in ways pre-proven in other aircraft, may not be fully represented or may be nonfunctional. Not all procedures can be performed dynamically, and as previously indicated, control-motion equations are not fully elaborated since the actual aircraft has not been tested and all the necessary inflight measurements and calculations are not completed.

Since, more and more, the current and expected changes in transport aircraft cockpits involve computer controlled systems and pilot information transmitted in digital form to multipurpose display surfaces, the prototype, stage d, simulator will include the software systems and proposed

advanced displays to the maximum extent that these developments are available. When actual components that conform to the published Technical Standard Order (TSO's) or specifications of the relevant industry standardization groups (particularly those of the Airlines Electronic Engineering Committee [AEEC] which is supported by Aeronautical Radio, Inc. [ARINC], and the airline industry) are not available, computers external to the stage d cockpit will be employed to simulate the relevant functions.

The final step in simulator evolution is, of course, the training level flight simulator. Since the intent in using this system is to substitute ground training for actual inflight experience, a maximum effort is expended to produce a duplication of actual flight deck capabilities and pilot stimuli. There are limits to the fidelity of motion and visual attachments, as there are to other environmental aspects; normally, an instructor position and a simulator operator position are provided, and the mere presence of these "others," as well as the records of performance being made, have the potential to affect attitudes and pilot behavior. Still, the experience of manufacturer and airline training departments provides conclusive evidence that a properly operated stage e, or training level simulator, provides an environment in which pilot behavior is highly predictive of actual flight experience. Crews that operate satisfactorily in the full performance simulator also pass flight checks, and special procedures that cause delays and difficulties in problem solution in the simulator are found to be similar to hard-to-correct problems in flight. The highest fidelity of simulation is required for training crewmembers for type ratings in a particular transport aircraft. These simulators provide exact replicas of flight crew stations, with operational equipment adequate to substitute for the actual aircraft in all significant aspects of flight operations and procedures. Two levels of simulation are typically used for training. The cockpit procedures trainer (CPT) shown in Figure 15 is a static device with operational discrete controls and displays having the capacity to simulate various normal, abnormal, and emergency conditions. Its purpose is to provide hands-on practice of procedures and task sequences and to illustrate what the results of correct and incorrect responses have on the aircraft displays. The motion base training simulator (Figure 16) is used to familiarize crews with the dynamic characteristics of the aircraft under normal flight conditions (including weather/turbulence variations) and, especially during emergencies that may be difficult or unsafe to simulate in the actual aircraft. The simulator is a six-degree-of-freedom device with a fully functioning replica of all crew stations in the cab, driven by a dedicated computer system capable of simulating the required flight conditions and aircraft characteristics. The visual system can be programmed to represent various airport/surrounding area configurations, under either night or daytime conditions and with varying visibility and ceilings. The training motion base simulator is used both for initial type rating training and for recurrency training per FAR's.

A stage e, training level simulator cannot be designed until the particular aircraft has been designed, and final completion of a "true" simulator requires data that are not developed until the reference aircraft has been built and flight tested; however, as inservice experience is accumulated, additional improvements in fidelity are made. Hence, a full performance simulator with these properties is never capable of utilization during the

flight deck design cycle. In actual fact, a prototype configuration simulator, representing stage d of simulator evolution, is never completed, with virtually all important systems, until the completion of flight deck design. Thus, the ground simulation activities accomplished during aircraft design are normally described as mockup studies (stage a), functional mockup tests (stage b), and limited examination of revised facilities in an engineering simulator (stage c).

3.5 Design Stage Flight Test and Simulation

As discussed in Section 2.3, flight tests and inflight simulations are conducted on a more or less continuing basis before and during the design cycle of any given new transport aircraft. A major new development such as an advanced navigation system, a change in cockpit flight displays from conventional electromechanical instruments to digitally controlled cathode ray tube (CRT) displays, or a computer added for flight management and fuel conservation functions, would never be injected directly into a new aircraft design without previous flight test and experimentation. Small changes in arrangement, simplifications in subsystems or improvements in flight deck procedures might be selected and examined *de novo* in connection with a particular design program, but the major or radical changes as illustrated, above, are too high-risk to be decided upon without closer replication of actual airline service experience than can be obtained in the laboratory, a mockup, or a part-task simulation.

Three examples of early design cycle flight test programs covering major flight deck changes may be cited. First, several developments in long-range navigation systems were pioneered by the military services and by international aviation organizations. Transport aircraft manufacturers were provided the results of several such test programs thereby making it feasible for them to install prototype advanced navigation systems in their own test aircraft. In some cases, NASA or FAA contracted with an airline or airlines to conduct flight tests of new over-ocean positioning systems, and again, the resulting data flowed to the manufacturers of transport aircraft.

As a second example of early flight test data utilization, digital-CRT flight instruments have been in flight test programs and have been installed in various military and civil test aircraft for five years or more. This flight test experience led to a decision by the domestic airline industry to support the choice of these systems for primary flight displays in new aircraft (see references 11, 12, and 13).

A third example of early design cycle use of flight test information is found in the case of flight management systems. Devices of this kind require no radical changes in the basic aircraft design, but rather tie together various already present flight data parameters with new computer processing for input to the autopilot and to primary flight displays.

Thus, it was feasible for both the avionics manufacturers and the airframe manufacturers to equip test aircraft with the new devices for extensive flight test programs. Due to competitive pressures, most such experimental work was conducted on a proprietary basis with the results leading to

design choices by the particular company involved. Published references to the design cycle test work were less often available than in the case of the earlier examples of early design phase flight tests conducted with broad industry or government participation.

Tests referred to in this section resulted in determining the acceptability of new concepts with particular concern for reliability, accuracy of presented information, adequacy of pilot performance employing the new system, and impact on normal and contingency flight desk procedures. Pilot workload, *per se*, and attempts to measure workload impact of the changes were generally of lesser initial interest so long as pilots reported no special problems.

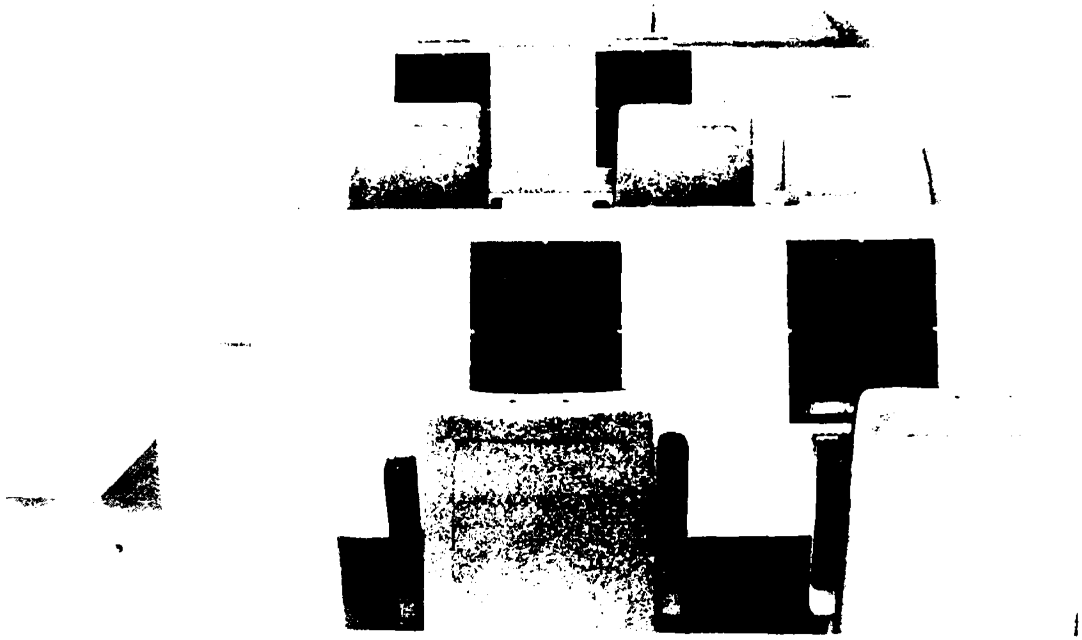
Flight test of major new concepts for cockpit systems and procedures can add important information to the design decision process. It cannot, however, substitute for the workload documentation program required under FAR 25.1523. One clear reason for this is that a number of individual changes and presumed improvements might be tested individually with each found to fulfill its promise, but this would not indicate the acceptability of the combined use. An example is the current proliferation of pilot input keyboards for navigation, communication, flight control, and flight management systems. There is no standard format, at present, and an increase in the number and diversity of keyboards may induce pilot errors or action delays.

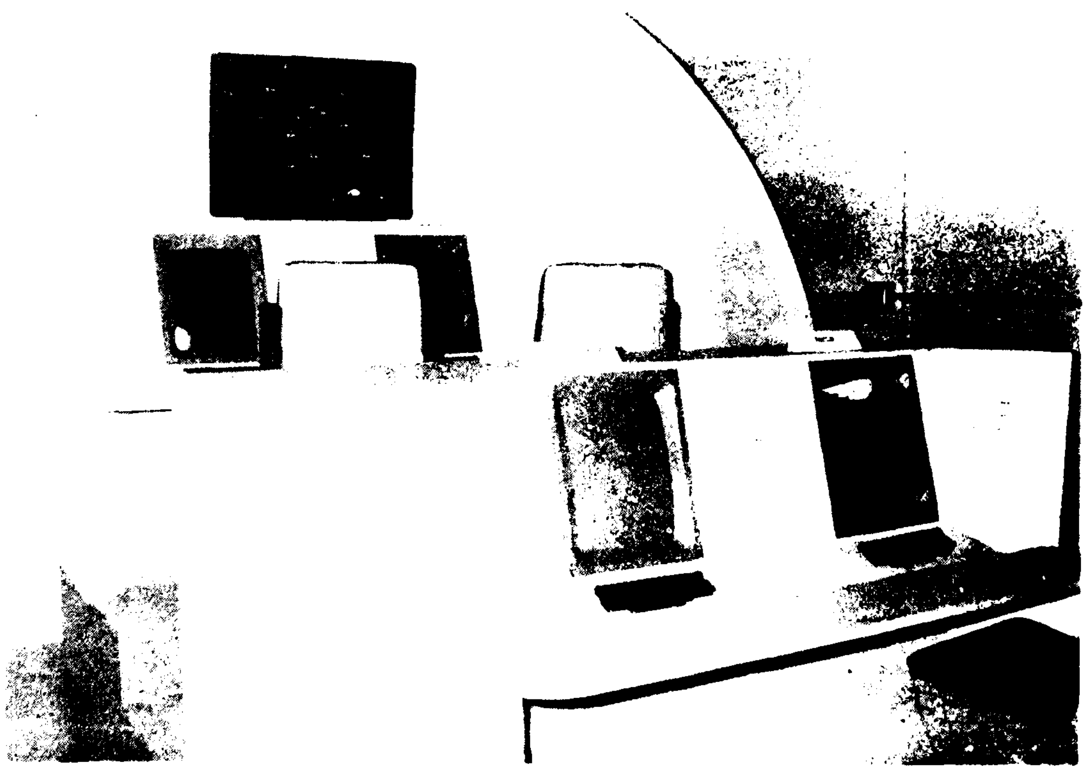
One of the most extensive programs of test flight using electronic flight instruments was the NASA-Boeing TCV program, previously cited. Alternate formats for displays were examined, reliability data were collected, and actual pilot performances in simulated landing were measured in a specially designed second cockpit of the TCV. This program provided the longterm, broad sample experience that is required to justify a major thrust toward revision of the conventional and service-proven prior flight deck arrangements.

No transport manufacturer simply adopted one or another of the configurations tested by NASA, or by the military in the Wright-Patterson Air Force Base digital avionics program, or by other agencies such as British Aerospace (who have conducted ground simulations with digital flight instruments) (see references 16, 17, 18 and 19). Rather than adopting some one earlier design, each manufacturer has proceeded to select concepts for change that appeared to be compatible with an overall flight deck configuration and interface them with related equipment such as a HUD or symbology and procedures made familiar by long use in selected flight director instruments. Hence, it is evident that pre-design flight tests are of great importance in proving the desirability of adopting some proposal for major change, but these flight tests do not substitute for, or replace, the detailed design process and workload test program.

Figure 4 Spatial Arrangement Mockup

LARGE SCREEN DISPLAY





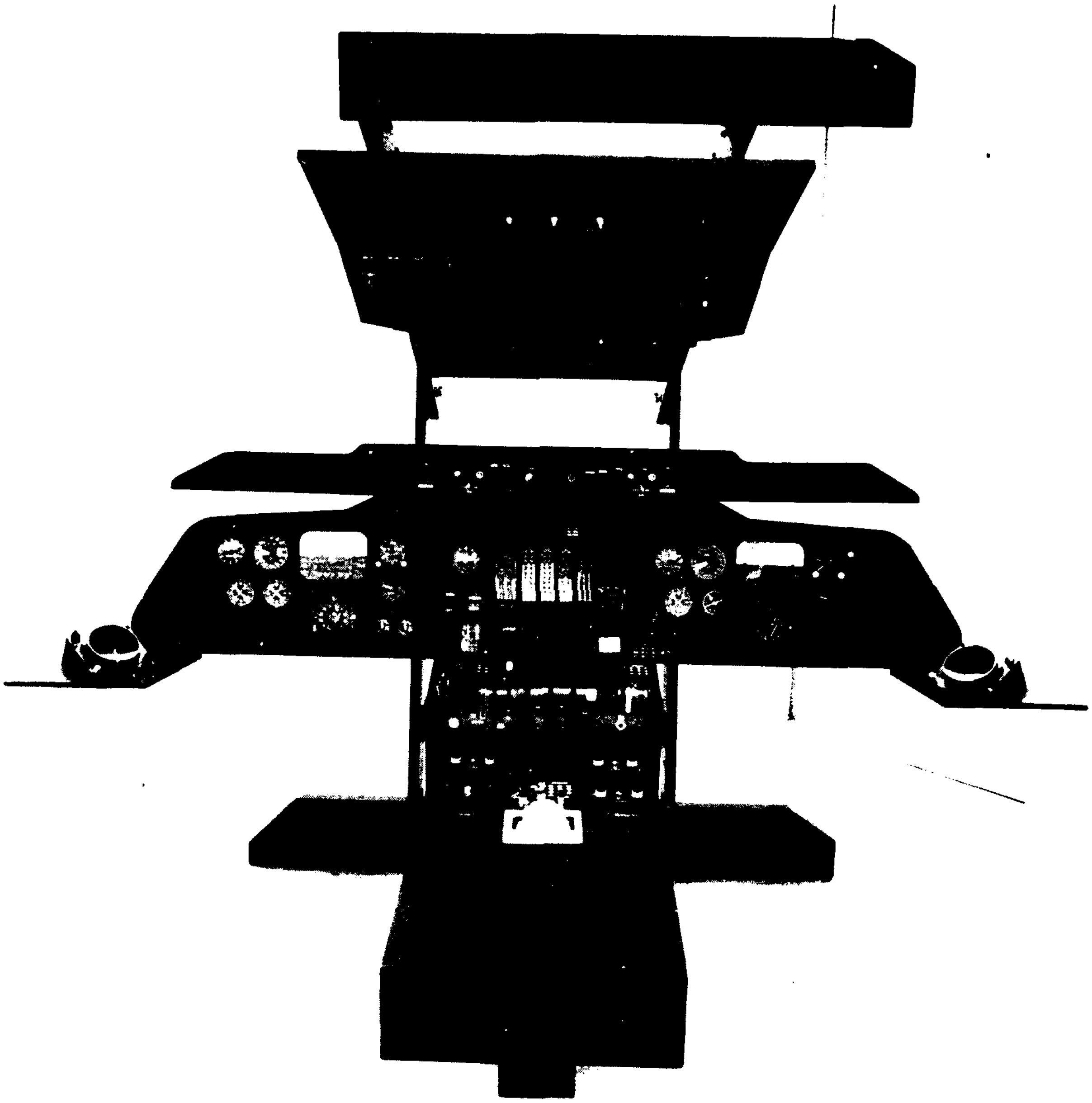


Figure 7 Static Mockup of Pilot & Copilot Panels

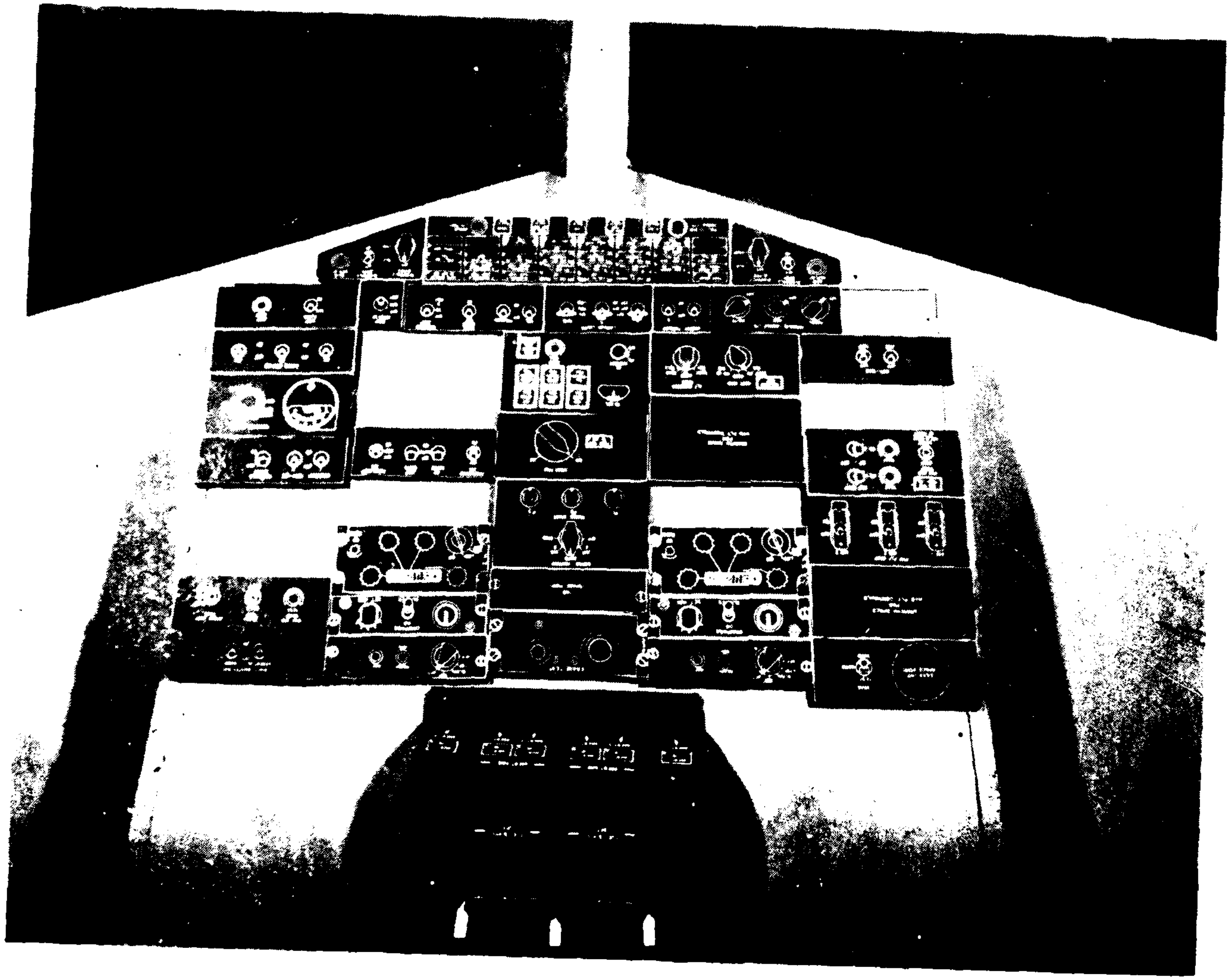


Figure 8 Static Mockup of Pilot & Copilot Overhead Panel

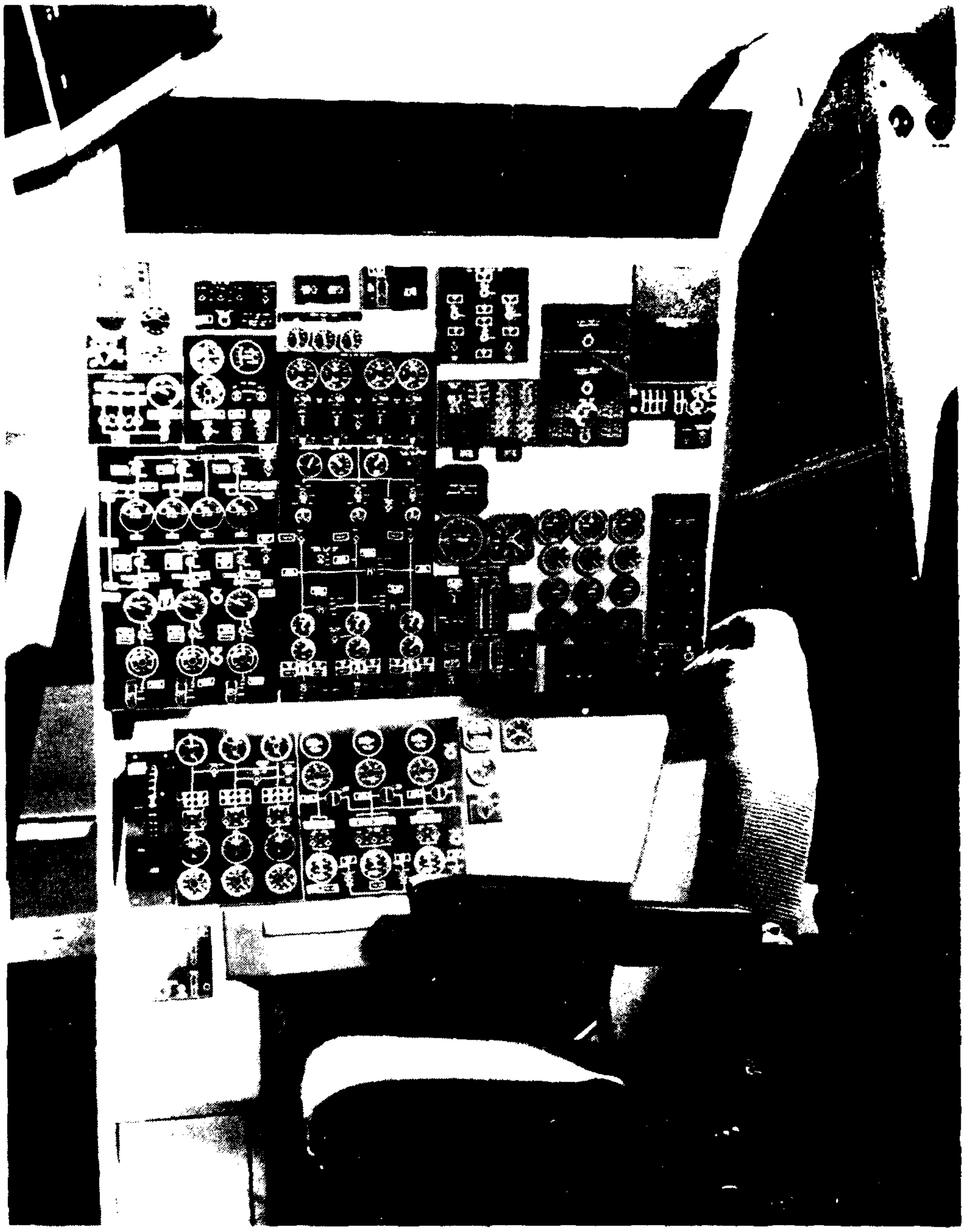


Figure 9 Static Mockup of Flight Engineer Panel

Figure 10 Flight Engineer's Station Lighting Mockup

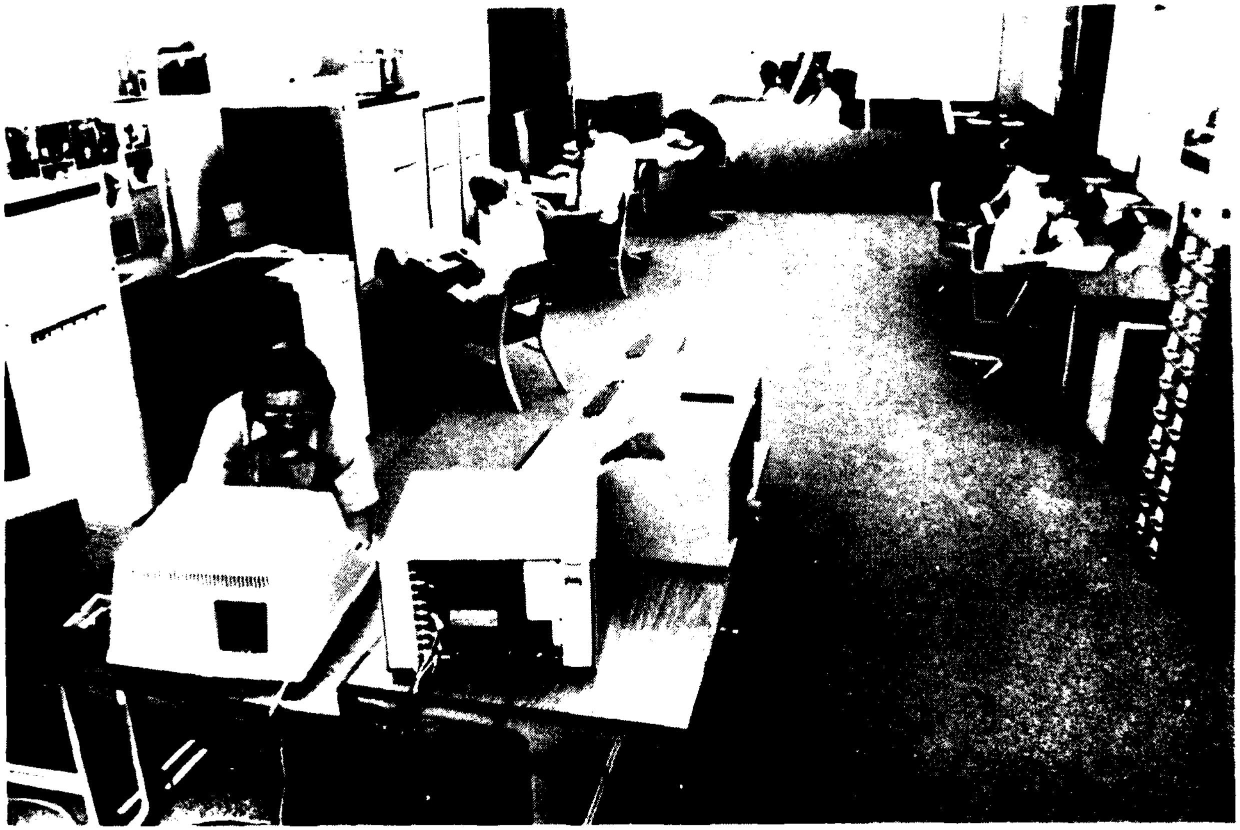


Figure 11 Douglas Digital Equipment and Technology Analysis Center

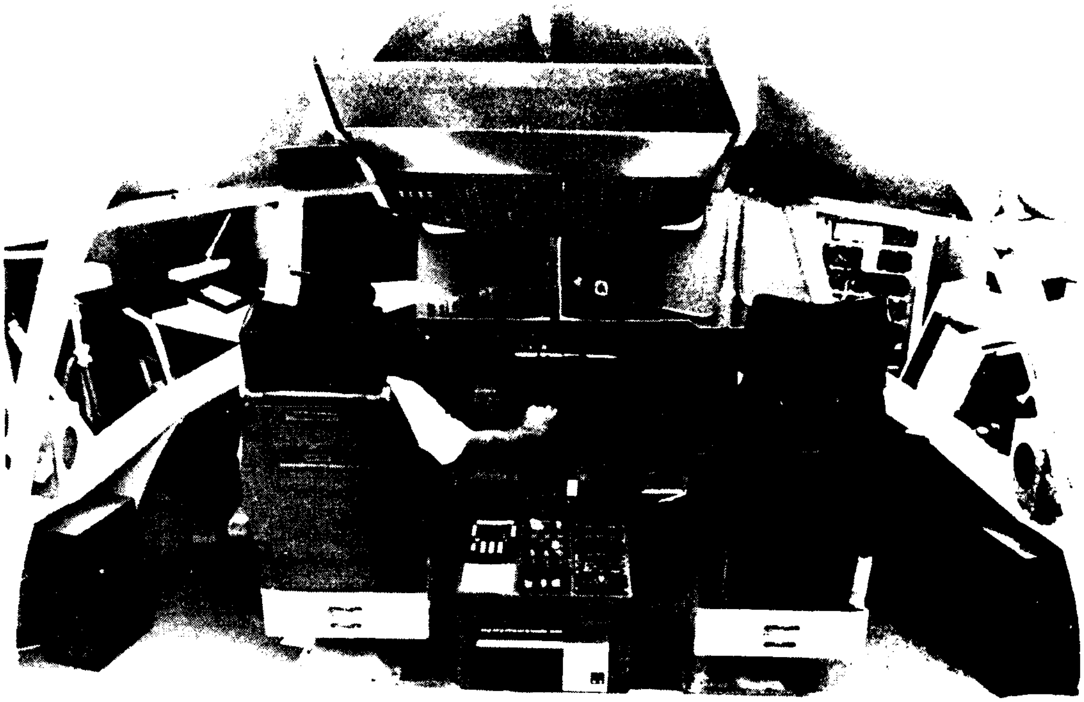
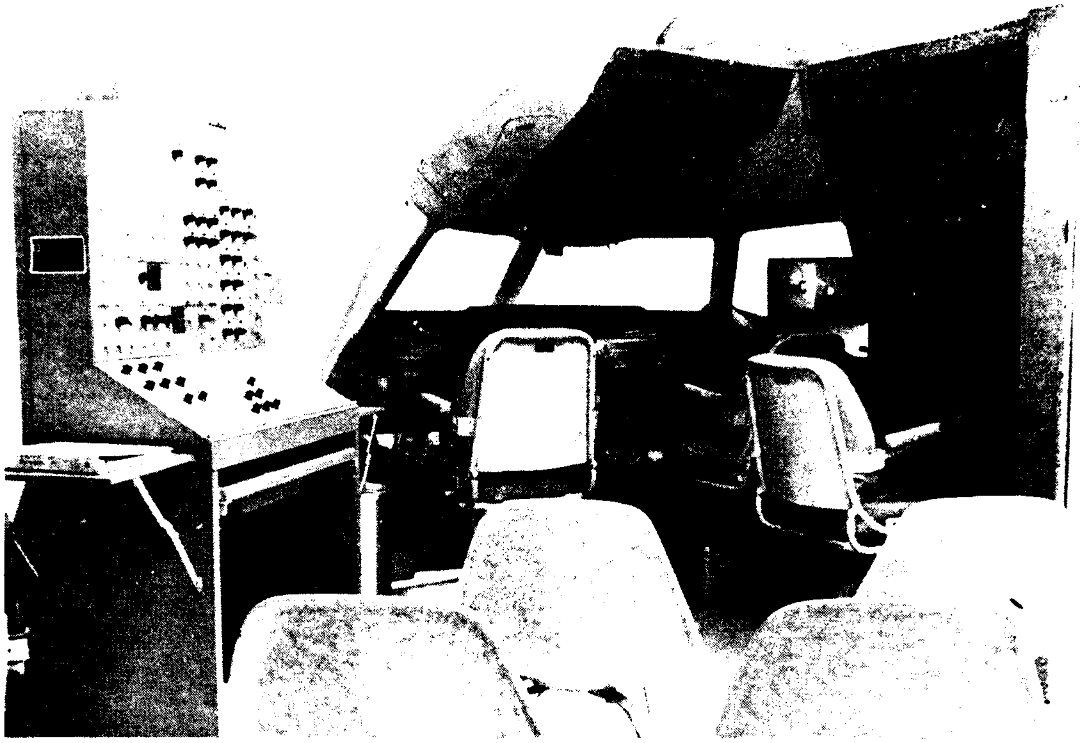
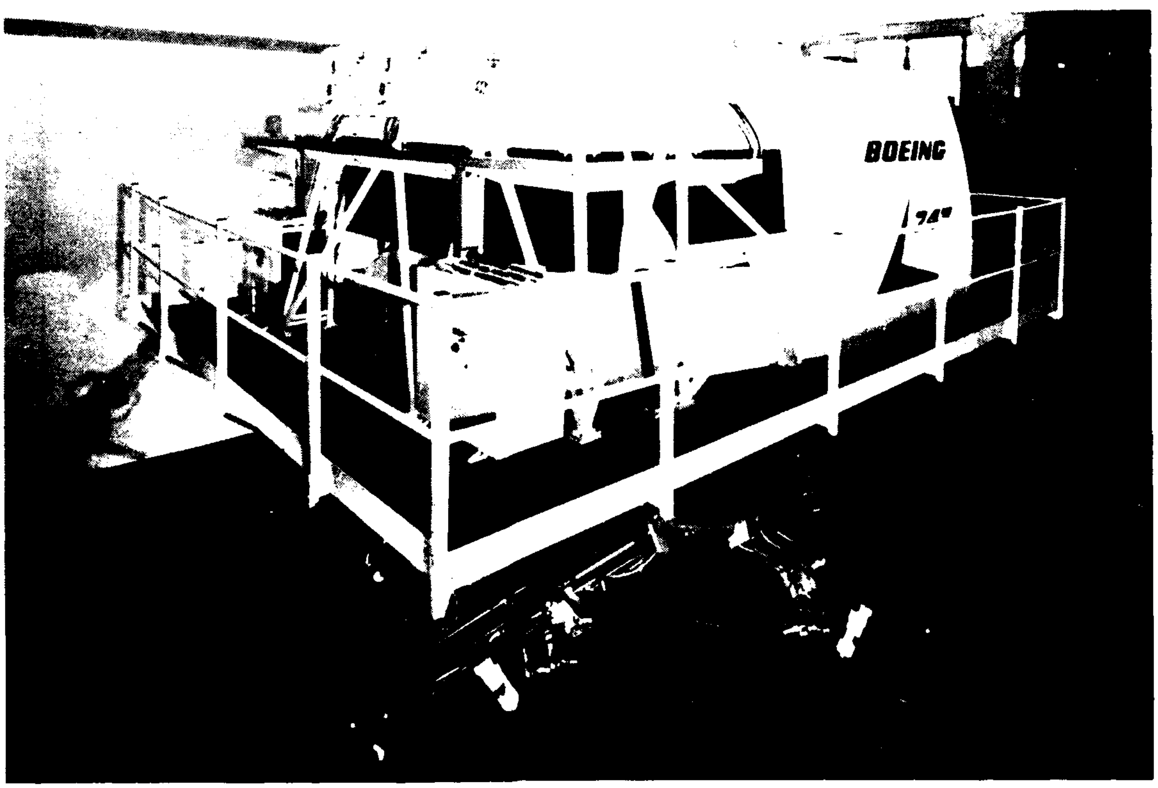


Figure 12 Cockpit Mockup with CRT-type Displays Incorporated

Figure 13 Cockpit Mockup with [illegible] Projection Screen

Figure 16 Typical Engineering Development Simulator




BOEING

4.0 Procedures After Aircraft Design

4.1 Update of Mockups, Advanced Studies

In one sense the aircraft design process never stops,and changes and improvements in flight deck facilities and procedures are made at various points extending throughout the actual production run of a model. Generally speaking, each aircraft order placed with a manufacturer is unique in some way or other going beyond the trim on minor appointments to the cabin. Recognizing this factor, the FAA requires the certification applicant to introduce a program for the individual certification of each individual aircraft, as made necessary by particular changes or special features. In Appendix A to this document, a sample manufacturer's application for single aircraft certification is illustrated, with differences described, flight deck panel changes portrayed, and a sample FAA "Aircraft Certification Eligibility" form appended. Still, it is convenient for the present purpose to note a milestone that represents the point at which design studies phase over to workload documentation studies with the beginning of planning for, and accumulation of, data intended to be submitted to the Type Certification Board to be considered in making the determination that the overall flight deck design and planned crew are "sufficient for safe operation."

At this program stage, the flight deck mockup will have been developed far beyond the initial pasting up of pictures and installation of controls from earlier models. The simulator increasingly will be furnished with actual equipment and controls conforming as closely as possible to those intended to be included in the production aircraft. Instruments will be lighted, many will be operable, and all system controls will be functional. As various avionics items become available from their manufacturers and suppliers, these and other flight deck components will be installed, so that the cab continues to progress toward the status of a true simulation of the final design.

It should be particularly noted that the initial design of a flight deck modification is based on previous work, normally with only a small number of planned innovations other than simplifications and other improvements that have been worked into previous production aircraft and shown to be entirely acceptable. For those systems that are being changed, design freezes are staggered by subsystems, often based on specifications provided to contract suppliers. When entering into contracts with suppliers of instruments, control devices, avionics and the like, an aircraft specification is issued as a part of each supply contract. After that formal action, the design of that portion of the flight deck is essentially frozen. A subsequent milestone occurs at the initial customer review when particular changes introducing differences in the overall configuration may be required. Obviously, the overall cockpit concept cannot be changed at customer option, except in the case of a major production order, but minor preferences or optimal improvements characteristically are selected. Because of the sequential nature of this process, it is not possible to compare all the new systems variations in a full-task simulation. The more usual procedure is to compare each new system or subsystem with a parallel system that is known to be acceptable on a prior aircraft.

Advanced studies will be conducted in the updated mockup as required to verify data flowing from computer model studies of task/motion and task/time lines. Availability of such an updated mockup makes it feasible to check, in a three-dimensional setting, any unusual or changed computations or predictions as to accessibility, visibility, and proper human engineering design.

In all aircraft programs there is a lighted mockup for use in checking on reflections, other visibility problems, general accommodations, seating, accessibility, etc. For use in this mockup, subsystem suppliers are required to provide a lighted, painted, dummy instrument early in the supply contract. It is not always the case, however, that this "stage a" simulation is developed through installation of functional systems and equipment. In aircraft programs with no major changes in configuration, there has been no requirement for advanced simulator experimentation.

4.2 Systems Analysis

Two of the preliminary methods of comparing crew workloads of different aircraft are by comparing check lists and procedures for those aircraft. Check lists are required for commercial aircraft by FAR Section 121.315 and must include "each item necessary for flight crew members to check for safety in engine and system emergencies." Check lists are just what they indicate, basic checks for safety items that are made for each mode of flight. Procedures, expanded procedures, or expanded check lists, as they are sometimes called, are more detailed explanations of crew requirements needed for operation of the aircraft.

Both check lists and procedures for normal operations, abnormal conditions, and emergencies can usually be found in the operations manual for each aircraft. An example of a check list and procedure item for gear retraction is:

> <u>Check List</u>
>
> Gear Handle - Up-latch check
>
> <u>Procedure</u>
>
> Raise gear, check/red/green lights, make up-latch check prior to 300 knots.

In the B-737 certification program, all check lists and procedures for the new aircraft were constructed and compared to three existing inservice aircraft, the other two-crew airline aircraft, the DC-9 and BAC 1-11, and the B-727. Except for the new aircraft, the procedures are taken from the manufacturer's operation manuals to ensure completeness. All of this material was organized under four headings: Normal check list (taxi check, before takeoff, after takeoff, descent and approach and landing); Procedures - system operation normal (fuel, miscellaneous, flight controls, power plant, hydraulics, electrical, APU, heat/anti-ice, etc.); Abnormal procedures (hydraulic pressure low, loss of system A, loss of system B, loss of both systems, pump overheat, etc); Emergency check lists (engine fire, engine overheat, tail compartment high temperature, etc.).

A count was then made to verify that the number of items was much less than that of the three-crew B-727 and comparable to the loading on the Captain and First Officer of each aircraft already in service. A count of this kind is not considered to be particularly useful since the extent and quality of workload varies too much from item to item. Only in-so-far as the items are highly comparable would a raw count represent a workload difference.

A major limitation of checklist comparisons is that these lists are focused on subsystem operation, the things that are done rarely, so they may be forgotten, unless a mechanized procedure of challenge and response is followed. Checklists do not cover the major pilot activities of flying the aircraft, dealing with air traffic control, and updating avionics for navigation. Much of pilot workload is omitted. The major attraction of checklist study is that it often is just in subsystem operation that workload reduction is sought (see Sections 2.4 and 3.2). Checklists are provided for important contingency conditions as well as for such normal operations at initiation of each phase of flight. To the extent that workload peaks are associated with emergencies, such as engine failure, evidence of reduction in the number and complexity of required checklist steps might be important.

Crew Procedures Objectives (CPO) represent expansion on the concept of checklists and are a later development from the Specific Behavioral Objectives (SBO) that were previously developed for use on training programs for the B-747 and which have been emphasized in recent years. CPO's are developed through a cooperative effort of the manufacturer and customer airlines. The principal purposes are to guide the training of pilots and to produce standardization that ensures that any qualified pilot on the company roster can function cooperatively and knowledgeably with any other company pilot as a disciplined team member. Thus, CPO's are not part of certification; the airline operating the particular aircraft can change them as appropriate. The activities related to development of an initial, industry-standard set of CPO's for a new aircraft are such, however, that a clearer understanding of task demands results. If there should be features in the new flight deck configuration that cause special workload problems, these features would be highlighted in the CPO documentation process and would be expected to come into general knowledge.

A minimum equipment list (MEL) comparison will usually be made as well. FAR 121.627(c) permits the publication of a MEL designed to provide operators with the authority to operate an airplane with certain items or components inoperative, provided an acceptable level of safety is maintained (see reference 20). The list does not include obvious required items such as wings, flaps, engines, etc., and it also may not include items unrelated to the airworthiness of the aircraft such as galley equipment and passenger convenience items. The list does include items related to airworthiness which may be inoperative for continued airplane operation, and the limitations and use of other operating components required when dispatching with inoperative equipment.

The MEL's for the new aircraft and for inservice aircraft may be divided into lists of items that cause different increments of inflight workload.

Major malfunctions are those requiring one or more of the following in-flight actions: (a) frequent monitoring, (b) different procedures during critical portions of flight, and (c) pilot having to leave his station in flight. Minor malfunctions are those requiring movement of switches or levers with subsequent occasional monitoring required. Insignificant malfunctions are those requiring movement of, at most, a few switches or levers with no subsequent monitoring.

A comparison of this kind was made when the B-737 was being developed. A count was made of the MEL items that might cause major, minor, insignificant, or no increases in workload for the new aircraft, with its several simplifications, and for a comparison aircraft, the B-727. The obtained information was discussed with particular consideration devoted to the questions: What role would a third crewmember have in dealing with inoperative equipment problems; What duties could he perform; and, What could he reach and see? In a somewhat similar discussion, for the initial model of the DC-9, an analysis was presented that enumerated the attained reductions in required procedures, estimated the workload impact of each advance in automation/simplification, and illustrated the limitations in external visibility from an added rear seat.

Historically, MEL has not been a required subject in crew complement certification. FAR 25.1523 and Appendix D do require examination of abnormal procedures and emergencies, but make no mention of MEL. As aircraft are designed to rely more and more on computer and digital systems, particularly in the case of relaxation of natural aerodynamic stability intended to reduce fuel consumption, it may be that the importance and criticality of MEL determinations is increased. The potential dependence of crew workload on the status of optional automation systems suggests that in the future it may be necessary to determine crew complement and the minimum equipment list in certification, and to do this with at least a "target" set of CPO's, showing the division of responsibilities and required actions among the individual crew members.

4.3. Task/Time Distribution

At this stage of the workload documentation program, there may be an updated developmental cab which is a realistic model of the planned final flight deck design, and there will be, in hand, final check lists and procedures for all major flight operations and contingency conditions. The next step is to utilize these products of the final design stages to generate actual data that predicts what the workload will be in the production aircraft, and more important still, how workload in the new aircraft will compare with that in a reference, line-proven aircraft. Normally, the goal of this test effort will be to show that workload is acceptable, a broad term that may cover several differences: a better balance of workload between crewmembers, reduction in peak workloads encountered in the reference aircraft, or a general reduction in subsystem monitoring and adjustment requirements, while actual flight control activities remain essentially standard. Since the general consensus among workload specialists holds that basic flight workload in current transport aircraft is acceptable, special simulation and task/time study activities performed after aircraft design may, alternatively, have as one of the objectives, the demonstration to the FAA Type Certification Board (TCB) that workload demands on the crew, and individual pilot workloads, are equivalent to, or

no more overall than, those in reference aircraft that have been proven acceptable in actual airline service. Comparable studies performed by the manufacturer prior to the essential completion of flight deck design were, of necessity, limited to part-task simulations. At this stage, however, there exists a complete enough description of the future flight deck to allow the manufacturer to initiate studies that approach the level of full-mission simulation. The data that are needed for such a demonstration of overall acceptability of the new design must be specific with respect to the crew seat, the type of flight operation, and perhaps most important, the comparison or reference aircraft. As discussed in Sections 2.4 and 3.3, the assumptions employed in arithmetic combinations of time durations for required actions for one analytical approach or time probabilities for another may be artificial depending upon the data base used by the particular manufacturer. Task time data from mockup studies, simulation and flight test may be used as well as data derived from time-motion techniques. Hence, the absolute numbers obtained do not predict real flight workloads; generally, they overestimate. For this reason, the obtained workload levels are useful primarily for comparison with similarly obtained numbers on a reference aircraft. Past experience shows that properly conducted synthetic studies can predict equivalence or the direction of a difference, more, or less, despite the limited applicability of the absolute numbers themselves.

Whether reported as a straight workload index reflecting percentage of available time occupied with crew-equipment tasks (see section 3.3) or in the alternative task/time probability distribution format (see section 2.4) the most important question is how should the results be interpreted. If the calculations result in peak workloads higher than those in critical flight scenarios in reference aircraft, it might be thought that the answer is simple--the design will not be acceptable. Several caveats must be stated on this point, however, since not all peaks are so high as to represent genuine problems, and there may be a compensatory gain. For example, in one DC-9-80 study, it was found that the first officer's workload increased in one phase of flight compared to the baseline aircraft (see reference 21). This peak increase was judged to be acceptable, since a better balance between captain's and first officer's workload was achieved, and in this particular instance, it was that balance, with reduced demand on the captain, that was most significant. There is, furthermore, some reason to emphasize average workloads, over reasonably brief periods, as having more meaning than essentially momentary peaks. The reasoning here is that despite CPO's and standard training, no two crewmembers will order required tasks in exactly the same sequence or with the same response latencies. Pilots recognize busy periods and adjust priorities and adopt task interleaving techniques to manage high momentary demands. Because of this built-in human capacity for flexible response, the data base task duration times, while appropriate for average situations of normal demand, overestimate the required task time under higher demand pressure.

Aside from a possible finding of peak workload in excess of that in the comparison aircraft under circumstances such as those just covered, the more usual finding has been that the resulting calculations for the new aircraft will show some workload reductions for particular segments and rough equivalence to reference aircraft in others. How, then, should this outcome be interpreted?

There is no general agreement that the best workload for aircraft crew members is 20 percent, 30 percent, or any other absolute figure. However, it has been observed that pilots start to drop less important tasks when 80% of their time has been occupied. The nature of the activities requiring allocation of a given percentage of available time, the difficulty of required mental efforts, and other considerations such as emotional stress, physical discomfort, professional pride and social satisfactions in a team-job well done, versus feelings of competition or hostility, will make a difference in any person's perception of how much work is enough and how much is too much. Hence, it is extremely difficult to decide what weight to place on obtained differences in a simple workload index.

In the face of these difficulties of interpretation, the nearly universally adopted solution is to compare measured or calculated workload during what are deemed to be the most critical minutes of a total cycle, usually the phase of final approach and landing. This is not to imply that workloads are not calculated for all phases of flight or that attention is not given to taxi, takeoff, climb, enroute, and descent phases. Rather, the point is that the most critical demonstration of workload suitability may be made for final approach and landing. The reasonable assumption is made that the individual crew members in proven inservice aircraft are capable of performing these critical operations -- as documented by the safe landing records of the comparison aircraft. Therefore, the task demands in those aircraft must not be too high, and any observer would probably verify that during approach and landing, they are not too low to sustain performance. If the measurements and calculations in the new design show that the individual crew members have equivalent or lower workload results during similar scenarios, it is generally concluded that the tests have shown workload to be acceptable. Equivalence is the key. Workload must compare favorably with another "good" airplane. Hence, there is no penalty, preventing the giving of credibility to the data, for lack of an absolute scale for optimum and obtained workloads. Comparison provides something more substantial than would such an arbitrary, and arguable scale. What this comparison standard is, really, is a summary fact based on many thousands of previous flights, accomplished by a wide diversity of pilots, under highly varied conditions of route, weather, traffic, and ATC. No known synthetic or laboratory procedure could possibly produce pilot performance data of this generality.

The outcome of such a comparison might be that approach and landing workload was slightly reduced, and this result would tend to verify the adequacy of the new design. But it could be the case that there were particular undesirable features to task demands that would show only in less busy flight regimes or in other unmeasured dimensions of workload. Due to the laboriousness of analytic procedures required to update the data base for task/time calculations, not all flight phases are covered intensively in all workload documentation programs. For example, the B-737 program looked at all flight phases but did not include a full analysis of enroute or "cruise" flight which made up about 60 percent of a typical scenario. Instead, equal segments were analyzed from takeoff, climb, cruise, descent, and approach segments. Furthermore, the nature of the measurements, time required or probability of being occupied, leaves out much that is important in flight deck evaluation. For example, an easy-to-perform, virtually

error-proof action may use equal time when compared to a more stressful action. "Quality" of work is only reflected in time required in an indirect sense, while "quantity" of work is specified. For these reasons, this form of measurement can fault a design, if it fails to show equivalence or improvement for the new aircraft, but it requires more development before it can be said to prove conclusively that the design is entirely adequate with respect to task demands.

The Douglas procedure developed for the DC-10 and used in measuring flight crew workload in the DC-9-80, is highlighted in Figures 17 through 21 (see reference 15). It is very much the same as the Boeing procedure. Both start with a typical flight plan developed to exercise all equipment of interest for this particular analysis. Using calculated or actual aircraft performance data and the hypothesized environmental factors, a mission scenario is developed along with an event timeline. The time available segments are then created by the time difference between events or milestones.

The operating procedures are established by coordinating inputs from experienced pilots using the published or proposed aircraft operating procedures for the proposed flight crew station configuration and the equipment operating procedures proposed for the planned configuration. These procedures are refined and detail times developed by using a full size design aid and published task time data to provide detailed task element time data. This establishes the time required for performing the various activities required for the flight.

The detailed flight sequences, time available data, and time required data are coded for acceptance by the computer program and entered into the computer for processing and analysis. The output from the computer consists of average and summary tabulations. These tabulations can be of total workload or the workload associated with various equipment and certain kinds of activities.

In a typical application of this task/time method, the procedures used reflect the operating techniques of one or two expert pilots who are experienced on similar aircraft. Other pilots might operate the aircraft slightly differently, but in a comparative analysis, since operating procedures are changed only when required by equipment changes, such differences are not important to the measuring system. The times calculated from the published data stores have been standardized by testing a large number of subjects.

This provides a solid base for use in identifying the changes in use times associated with change in the crew station configuration. It is this ability to identify change in use time as it affects overall workload that is most useful to comparison of one design with a reference design such as that employed in an in-service aircraft.

In the Douglas procedure, a separate routine referred to as the Select Option allows retrieval of task elements in accordance with specific ends associated with equipment, body actions, reach distances, responsible engineering groups, etc. Workloads are then computed for these groupings. By proper selection, it is then possible to determine which factors are significant in relation to overt workload or where the configuration could

be improved more effectively to reduce workload peaks. This technique enables comparisons to be made on a numerical basis independent of personal opinion.

Developing the Time Required starts with the use of flight crew station drawings, proposed operating procedures for the aircraft, and operating procedures for the specific equipments proposed in the configuration under consideration. In close coordination with pilots experienced in similar aircraft, a very detailed description of the procedures required for flying each mission segment is developed. Figure 18 gives a graphic portrayal of the last 15 minutes to touchdown, and Figure 19 is an example of the worksheet used in developing the procedures, time required, designating the crew member, and indicating the control or display involved. The time required for each action is calculated by using a standard data store that provides the times for actuating various types of controls and reading various types of instruments (see reference 14).

Douglas data are organized in terms of missions, functions, tasks, and elements. Symbols and a brief description for each of these are used to facilitate both the computer processing of the data and the reading of computer printouts. Figure 20 simulates a computer printout for some of the elements in the landing phase. Some of the items are as follows.

The computer program provides the capability to summarize and average the workload between designated milestones as related to all task elements or those selected on the basis of equipment or specialized task requirements. It shows the flight crew workload summary for all task elements in a portion of the flight. The computer prints out the function symbol, task symbol, task description, and the percent workload for the Captain and the First Officer. The minus sign in front of the workload numbers is due to the technique of counting time back from the touchdown point.

Figure 21 shows a typical workload breakdown by equipment groupings and activities. It shows, for example, the high proportion of workload associated with the communications task in a high density area such as New York. The low value of workload required for outside scanning corresponds to the IFR conditions chosen to give high workload in other areas. The total level of workload, however, would allow outside scanning to be increased considerably if needed. Outside scanning is also possible simultaneously with some of the other task elements listed. Figure 22 shows the detailed analysis that is possible when the workload elements are displayed for each task. This shows the major elements of the workload structure and how they change from task to task. Most of this 5 minute segment is with the aircraft on autopilot, and the Captain takes over manually for the last 200 feet.

Though the Boeing method is similar to the Douglas method, Boeing's data summaries are organized differently. Timelines and workload profiles for the various channels are developed for workload overviews and for more detailed analysis as presented in Figures 23 through 37. The resulting workload profiles give the workload for each channel of concern, with percent workload on the ordinate versus lapsed time on the abscissa. The time history presentation method permits a quick scan to determine if any undesirable work loading conditions exist. If the workload time history profile is

within pre-established bounds, results demonstrate an acceptable workload level. If not, the peaks can be examined more closely compared to timeline and subsystem activity data, to determine which tasks and which subsystems contribute to the high levels. Multiple user options exist to facilitate this process. From such reviews, the analyst can then reappraise the procedure, task and subsystem design to decide what type of design change would most satisfactorily alleviate the high load condition.

As stated previously, these workload evaluation methods parallel design for trade-off purposes and permit comparative evaluation of alternative crew station configurations, the new aircraft flight deck and that of comparison in-service aircraft.

4.4. Simulator Studies

There is an iterative process in the overall development of the data base used to generate computer calculated workloads for new and reference aircraft. Task sequences originally generated by pilots, see Sections 3.1 and 4.2, are reviewed by pilots and updated accordingly. Then, flight crew task sequences are conducted in the mockup or developmental cab. These simulated flight segments may be video taped for subsequent review to validate the tabulations. In this way, qualified pilots are employed to check and verify the subtask definitions. Further, times associated with any new features and procedures in the new flight deck are defined and entered into the data base. Changes may be made in the sequencing of tasks and many revisions made to obtain as representative a set of SOP's and associated task times as possible. This includes updating of sequences to reflect optimum crew techniques for utilization of new equipment and system capabilities in the new design. Typically, there may be as many as eight iterations of the flight scenarios and over eight hundred tasks identified separately for each crew member. The flight deck mockup serves for verification data collection in each iterative cycle. In this way, maximum assurance is developed that the task data base is complete and accurate for comparing crew workloads in the new and reference aircraft.

4.5 Visibility and Collision Avoidance Studies

In the early stages of a new aircraft program, long before an actual workload documentation effort is initiated, it is usual to use a special computer program to estimate the effectiveness of flight deck window design. Typically, a subprogram will calculate window size in steradian measures for each observer's field of vision and intersecting fields of vision. With these results it is convenient to generate a plot of control cabin windows in which the boundary of windows is shown from each crew member's design eye position. One of the most useful forms of such a plot has been produced in color so that the fields of view from each eye position and areas of overlap are chromatically portrayed.

For the workload documentation program, the results of these early design stage efforts may be appended in the form of a comparison between the amount of outside visual field available for the crew members of the new aircraft versus the reference aircraft. If there is a difference in crew size, the contribution of an observer will be evaluated.

The relationship between available external vision and collision avoidance may be computed for different speeds and crew members. The size of the visual field is the primary factor in collision avoidance but not necessarily in traffic detection. This is because of the differences in detection of traffic on a collision course and traffic not on a collision course. Traffic that is on a collision course: (1) will almost always remain at the same relative angle or appear to remain stationary in the window, and (2) will appear progressively larger. Whereas an aircraft not on a collision course may normally appear to move (have relative motion), and its image may become larger, smaller, or remain the same.

The human eye has only a very small field of high acuity. This property is a function of the distribution of the rods and cones of the retina and gives the eye its directional sensing ability.

For example, a 12 foot diameter sphere at a distance of 55,000 feet will subtend an angle of 3/4 of a minute of arc at the eye. For the eye to detect the sphere at this distance, it must be aligned so that the field of acuity for 3/4 of a minute of arc includes the sphere. For a pilot to scan his total visual field of 2.99 steradians for an object of this apparent minuteness at that distance would require 20 million different eye fixations.

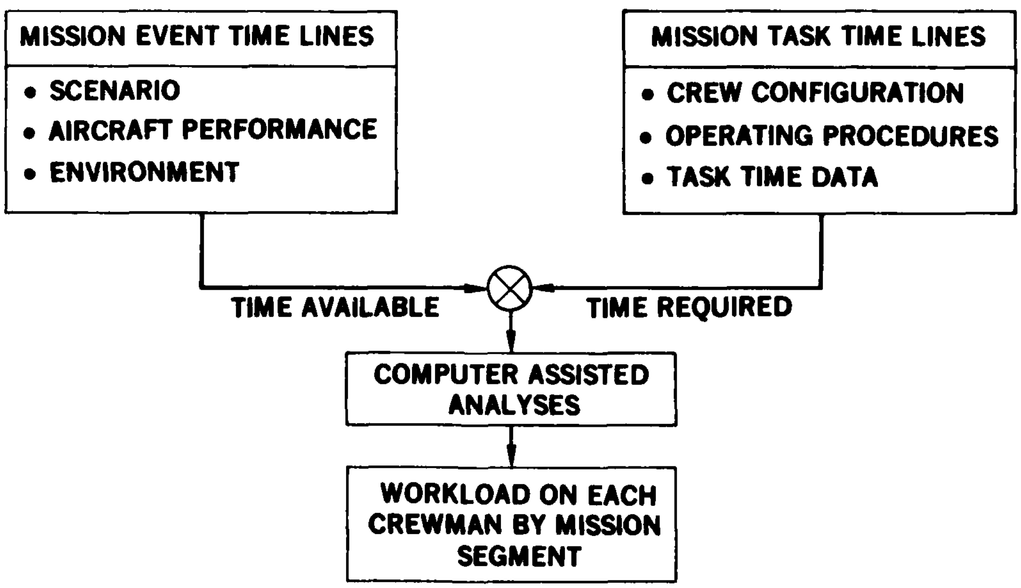


Figure 17 Workload Analysis Methodology

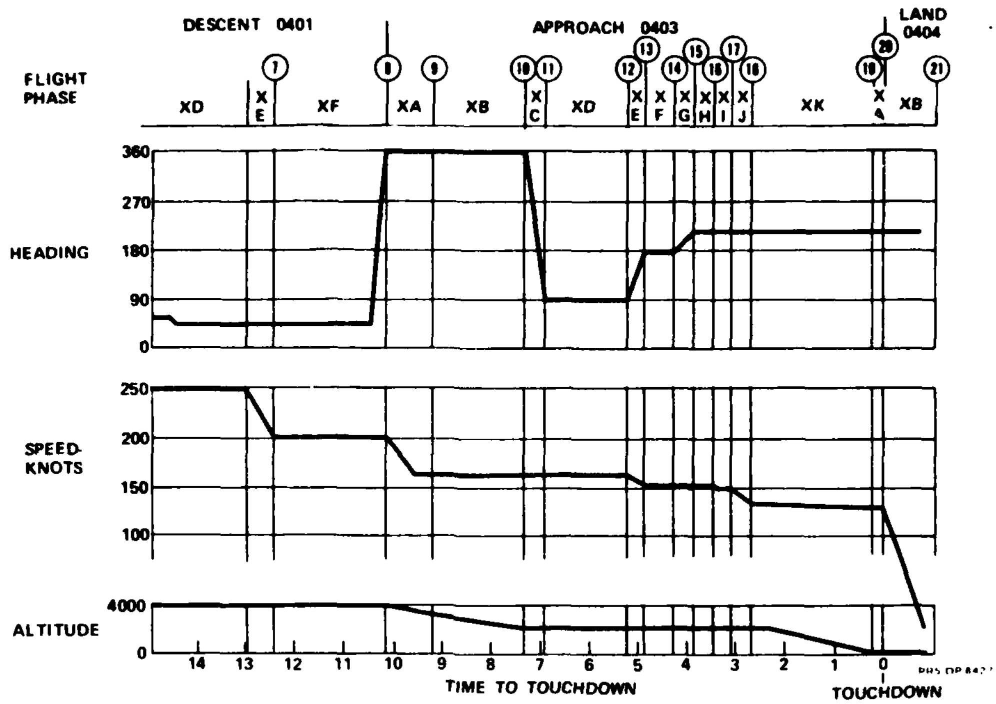


Figure 18 Flight Profile Segment

CHANGE NOTATION 00 PAGE INDEX ______ ADDITION NO. ______

FLIGHT PHASE IDENTIFIER 0403 TITLE APPROACH

TASK IDENTIFIER BL TITLE SET FLAPS TO 50 DEGREES

NORMAL X
CONTINGENCY ______
DIL ______

CONFIGURATION ______
EQUIPMENT INTERFACE ______

CRITICALITY ______
DEFERABILITY ______

MISSION TIME ___ : ___ : ___ HRS : MIN : SEC

ELAPSED TIME ___ : ___ : ___ HRS : MIN : SEC

TOTAL TASK TIME REQUIRED ______ SEC.

ELEMENT	DESCRIPTION	T_R	C	FO	C&D REFERENCE	REMARKS
01	CAPTAIN CALL OUT: FLAPS 50 DEGREES	1.0	X	X		
02	GRASP FLAP/SLAT HANDLE AND DEPRESS AND HOLD UNLOCK LEVER	1.4		X	CP-FL	
03	PULL FLAP CONTROL HANDLE TO 50 DEGREE DETENT POSITION	1.9		X	CP-FL	
04	RELEASE UNLOCK LEVER	0.6		X	CP-FL	
05	MONITOR FLAP POSITION INDICATOR UNTIL IT READS 50 DEGREES	8.0		X	PC-55	
06	FIRST OFFICER CALL OUT: FLAPS 50 DEGREES	1.0	X	X		
07						
08						

Figure 19 Typical Flightcrew Workload Measurement Worksheet

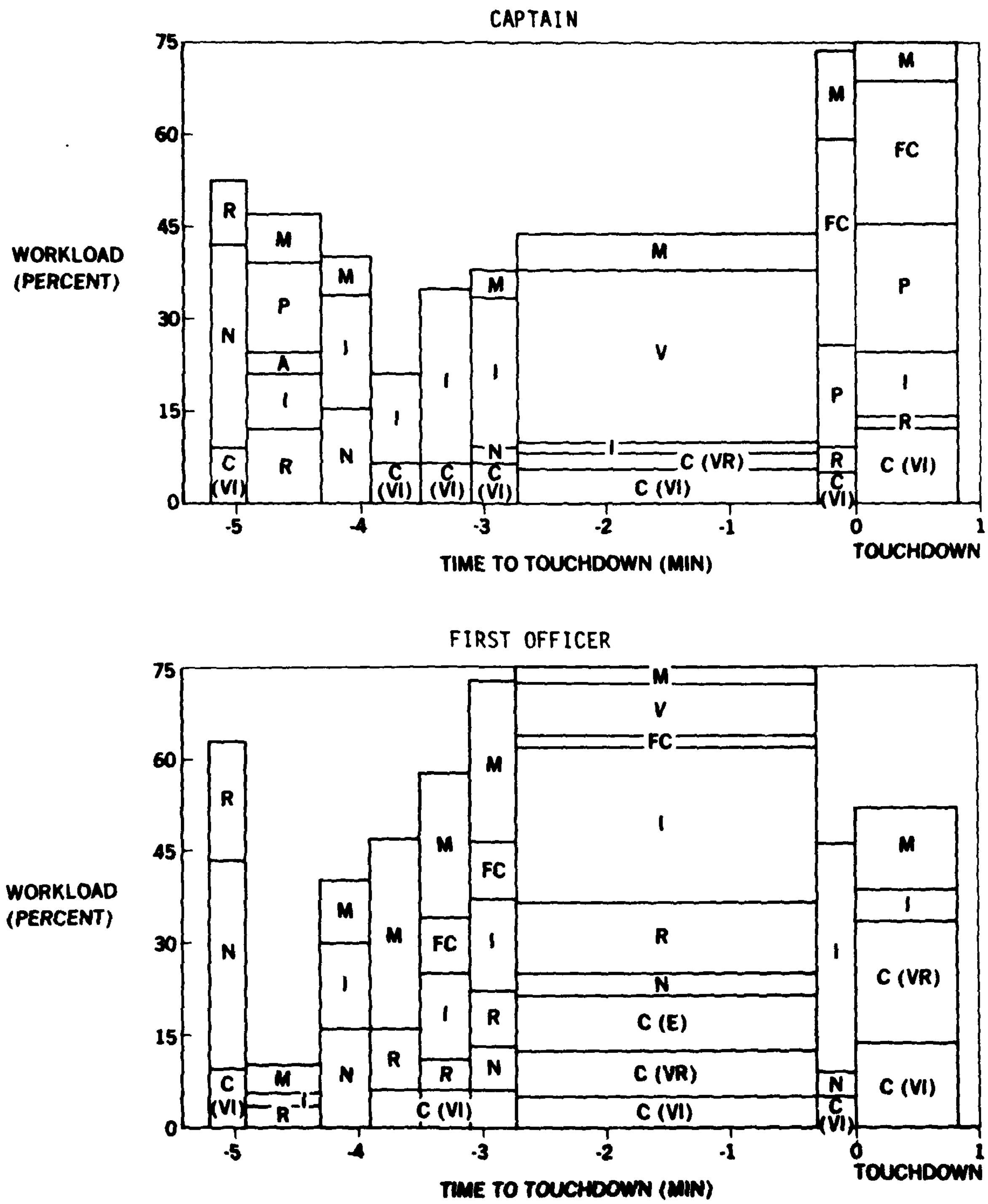
WORKLOADS ARE FOR ALL TASK ELEMENTS

FUNC	TASK	TITLE	TIME H M S	DURTN MIN	PERCENT WORKLOAD C	FO
0401	XA	COMMENCE DESCENT TO 10,000 FT; REDUCE SPEED	00:20:18	2.00	-77.92	-72.75
0401	XB	END DESCENT AT 1500 FPM;	00:18:18	3.00	-34.22	-28.72
0401	XC	END DESCENT TO 4000 FT	00:15:18	0.80	-4.17	-31.88
0401	XD	COURSE CHANGE TO HDG 044	00:14:30	1.60	-89.17	-82.29
0401	XE	START SPEED CHG 250 - 200 KIAS	00:12:54	0.50	-78.67	-63.67
0401	XF	END SPEED CHANGE	00:12:24	2.30	-27.61	-67.61
0402	XA	THIS PHASE NOT USED	00:10:06	-0.02	0.00	0.00

Figure 20 Flightcrew Workload Summary (Simulated Computer Printout)

EQUIPMENT	FLIGHT PHASE WORKLOAD DESCENT CAPT	F.O.
TOTAL COMMUNICATIONS	19.5	32.9
VERBAL EXTERNAL	(10.3)	(16.4)
VERBAL INTERNAL	(8.9)	(8.9)
EQUIPMENT	(0.3)	(7.6)
FLIGHT INSTRUMENTS	10.0	3.0
NAVIGATION INSTRUMENTS	7.5	7.0
REACH	3.8	9.0
SYSTEM SCAN AND CHECKLIST	0.0	6.5
POWER PLANT	7.4	0.8
FLIGHT CONTROLS	0.0	0.0
AUTOFLIGHT	1.4	0.2
MISCELLANEOUS	3.2	1.4
NORMAL OUT SCAN (IFR APPROACH)	4.9	4.9
TOTAL	57.7	66.7

Figure 21 Examples of Crew Workload by Equipment or Activity



M = MISCELLANEOUS
FC = FLIGHT CONTROLS
V = VISUAL SCAN
P = POWERPLANT
A = AUTOFLIGHT
I = FLIGHT INSTRUMENTS
R = REACHES
N = NAVIGATION
C (E) = COMMUNICATIONS EQUIPMENT
C (VR) = VERBAL COMMUNICATIONS (RADIO)
C (VI) = VERBAL COMMUNICATIONS (INTERNAL)

Figure 22 Flightcrew Workload Structure Analysis

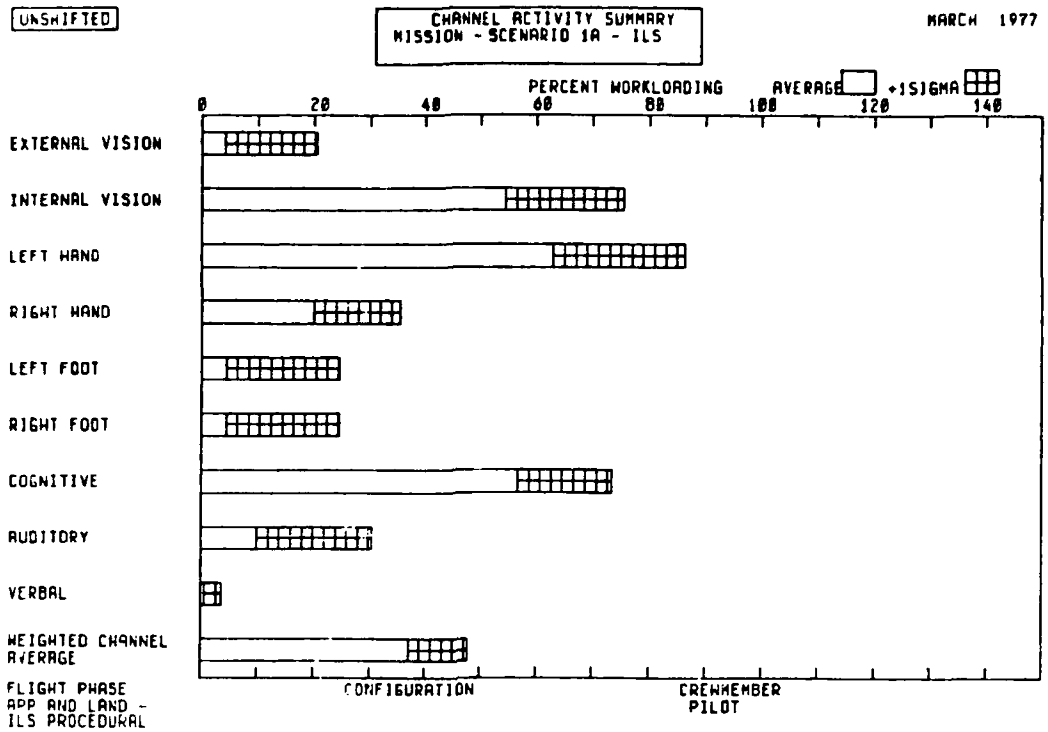


Figure 23 *Channel Activity Plus One Sigma Bar Chart Summary*

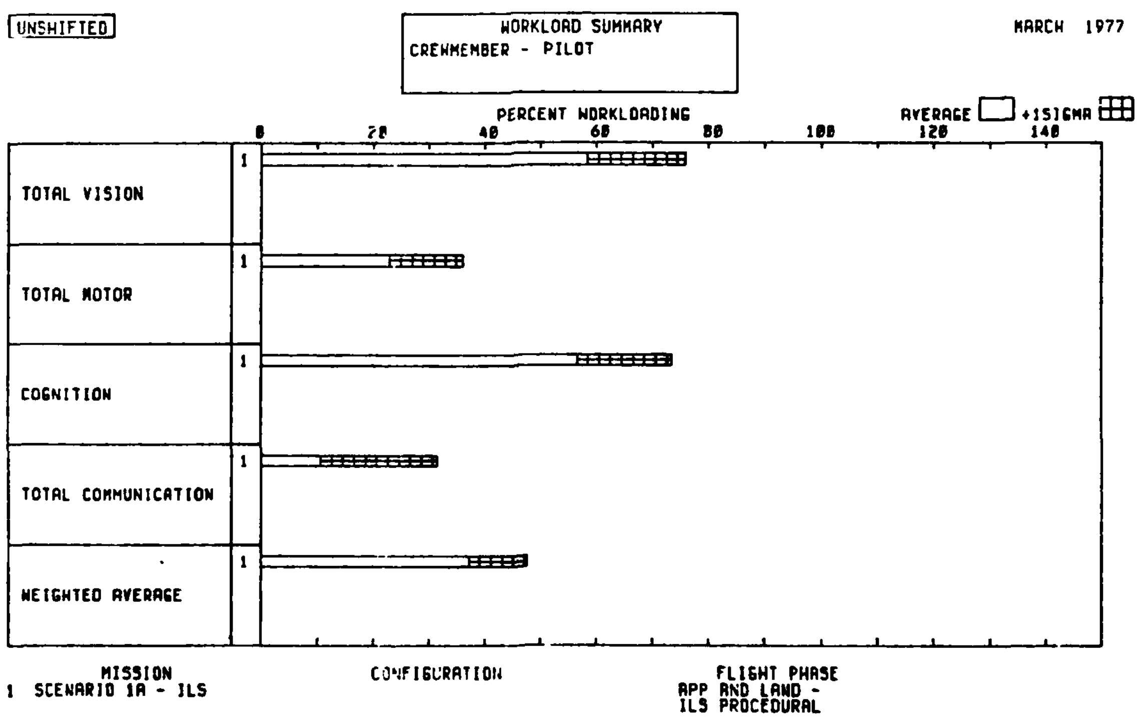


Figure 24 *Workload Barchart Summary Plus One Sigma*

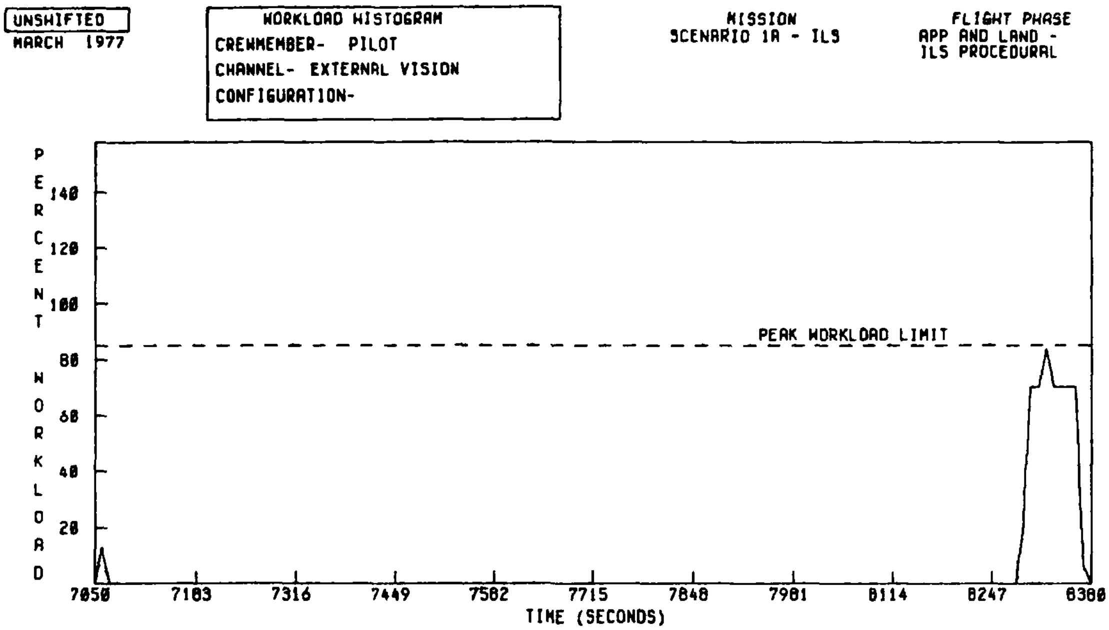


Figure 25 *External Vision Workload Time History*

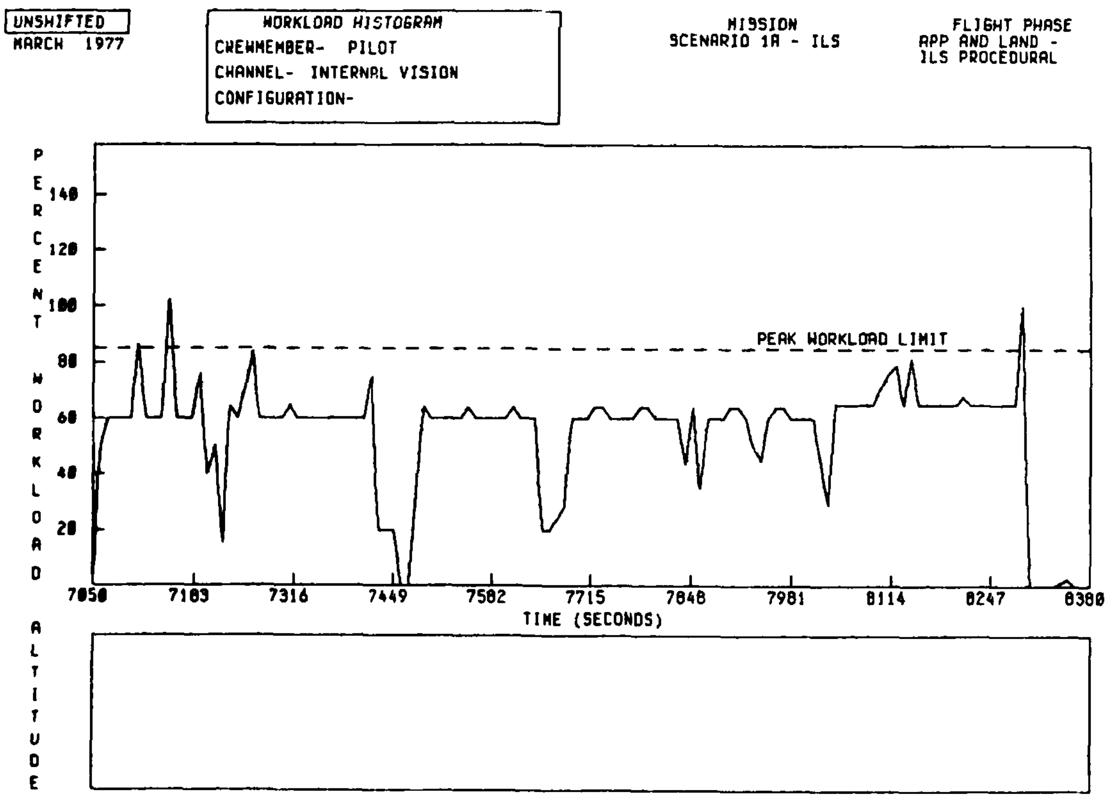


Figure 26 *Internal Vision Workload Time History*

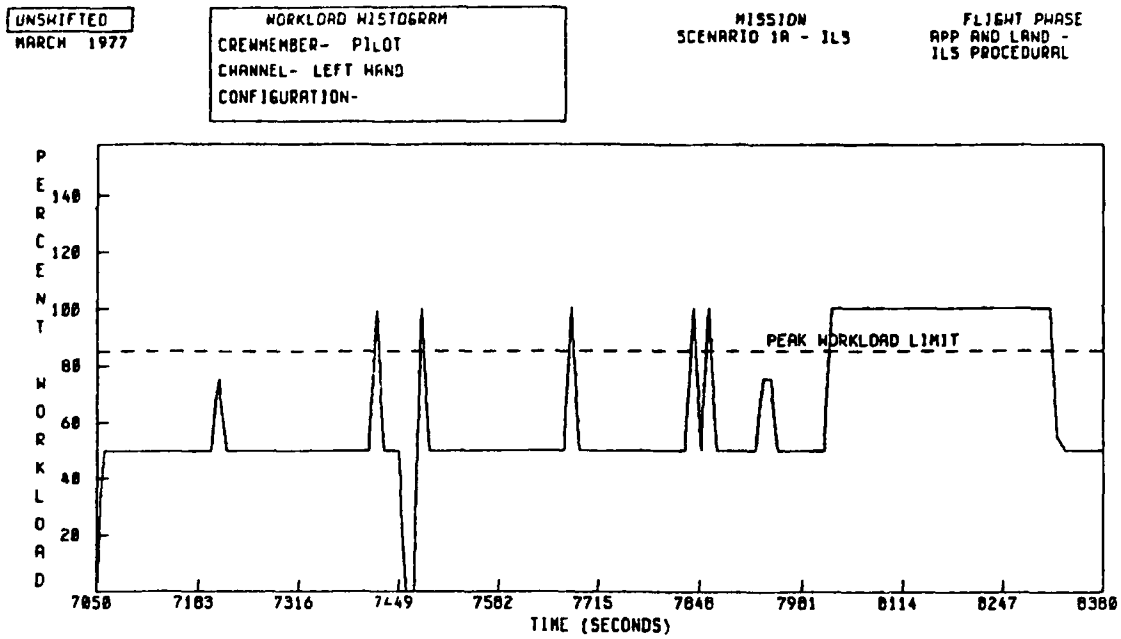


Figure 27 *Left Hand Workload Time History*

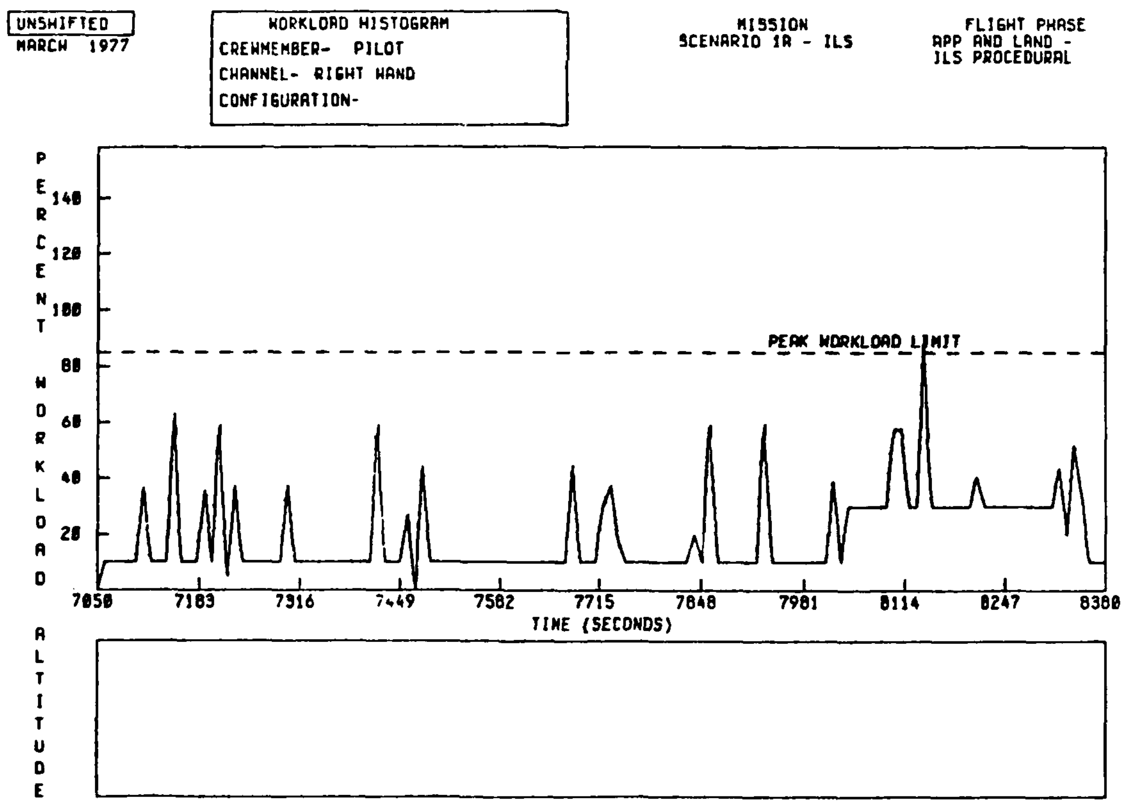


Figure 28 *Right Hand Workload Time History*

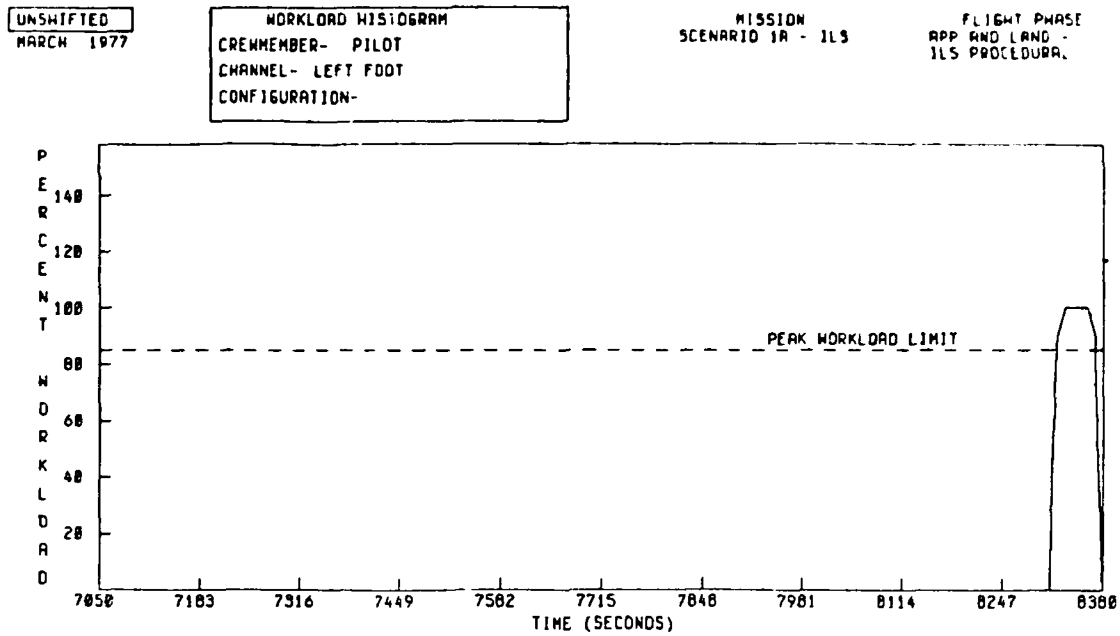


Figure 29 *Left Foot Workload Time History*

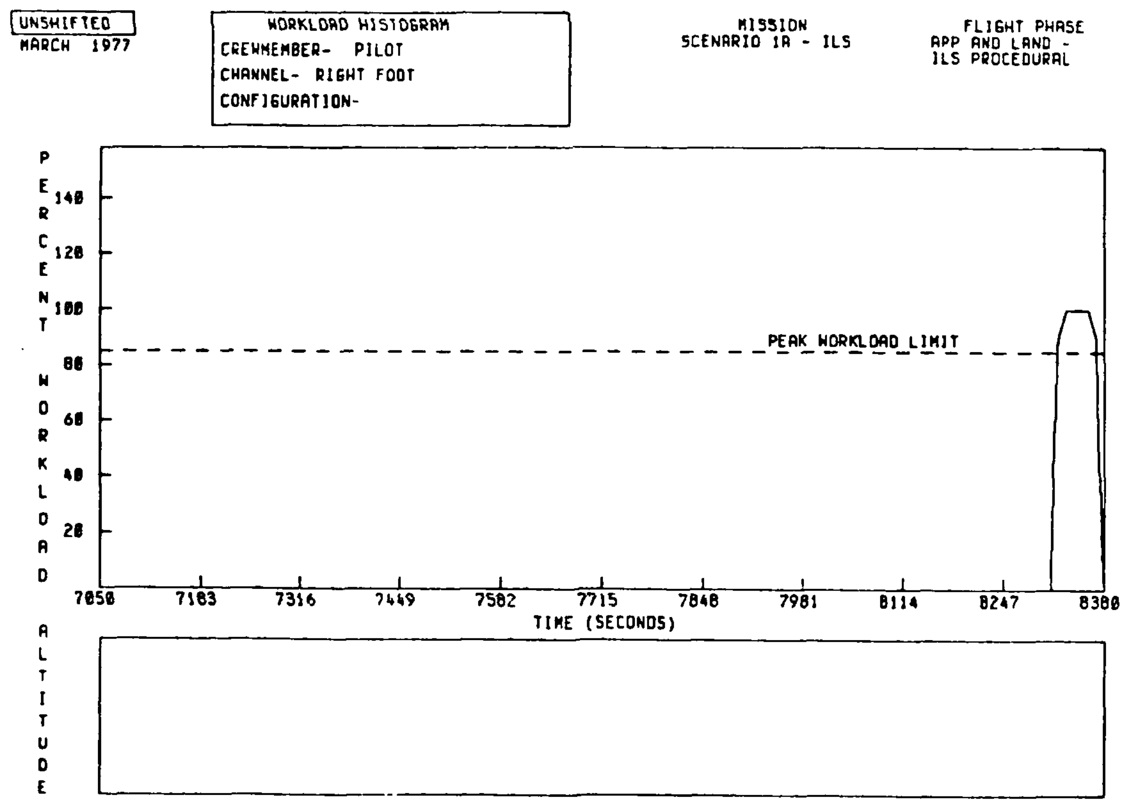


Figure 30 *Right Foot Workload Time History*

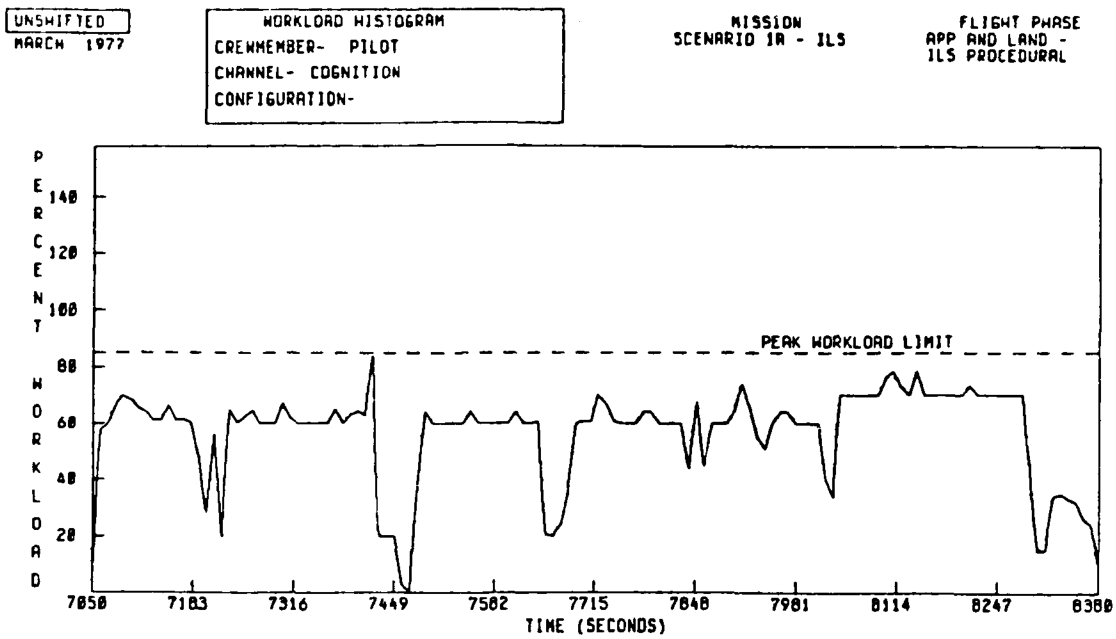


Figure 31 *Cognitive Workload Time History*

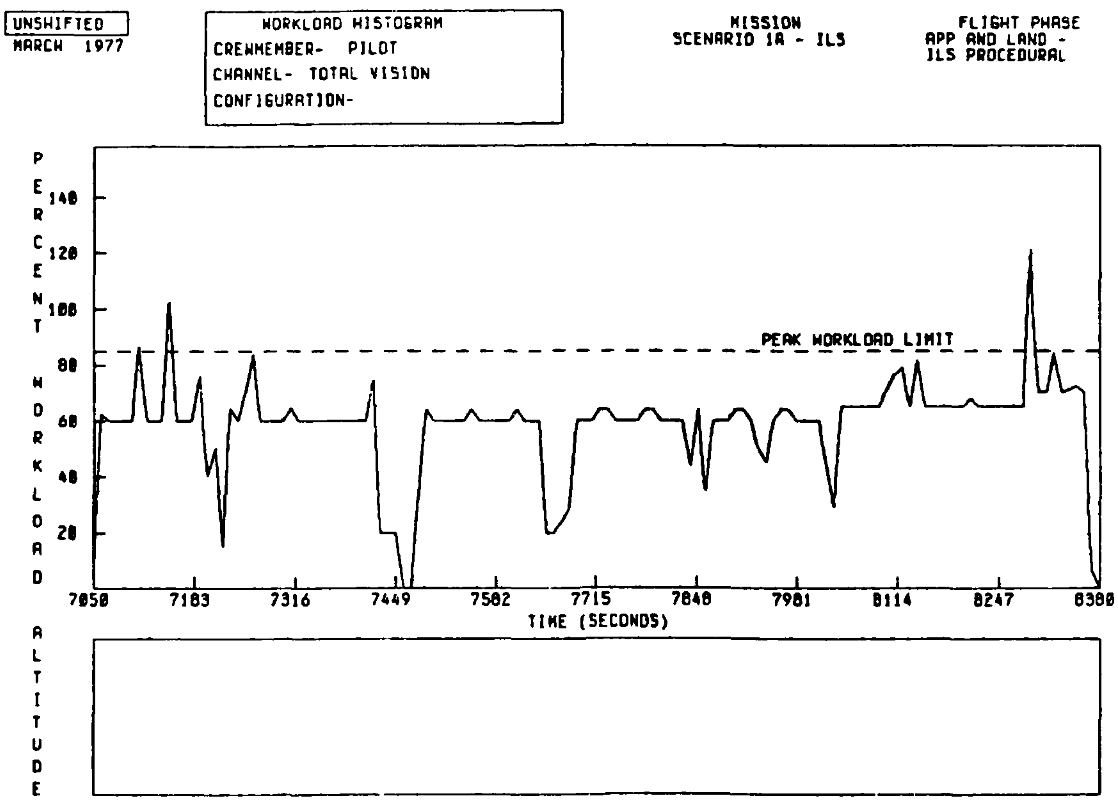


Figure 32 *Total Vision Workload Time History*

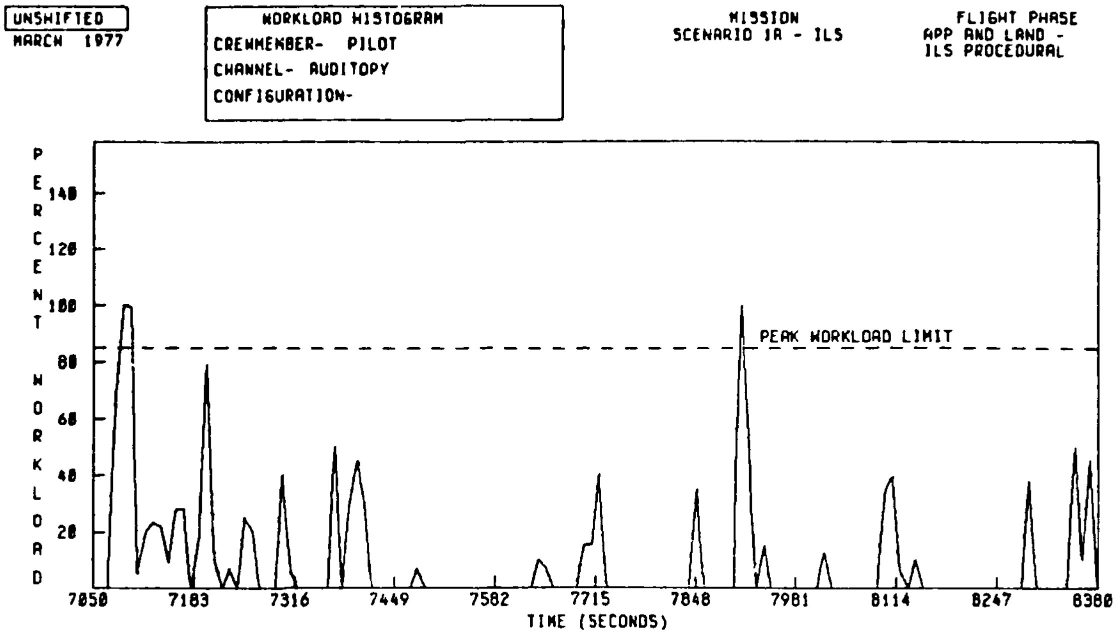


Figure 33 *Auditory Workload Time History*

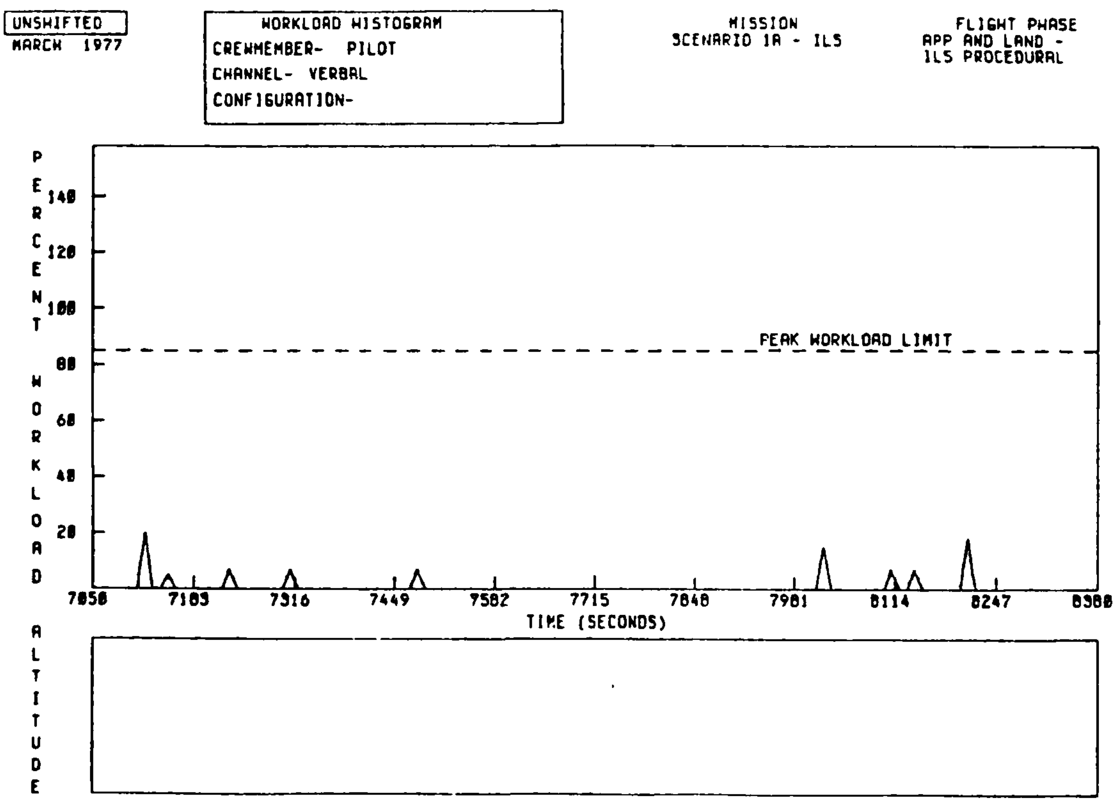


Figure 34 *Verbal Workload Time History*

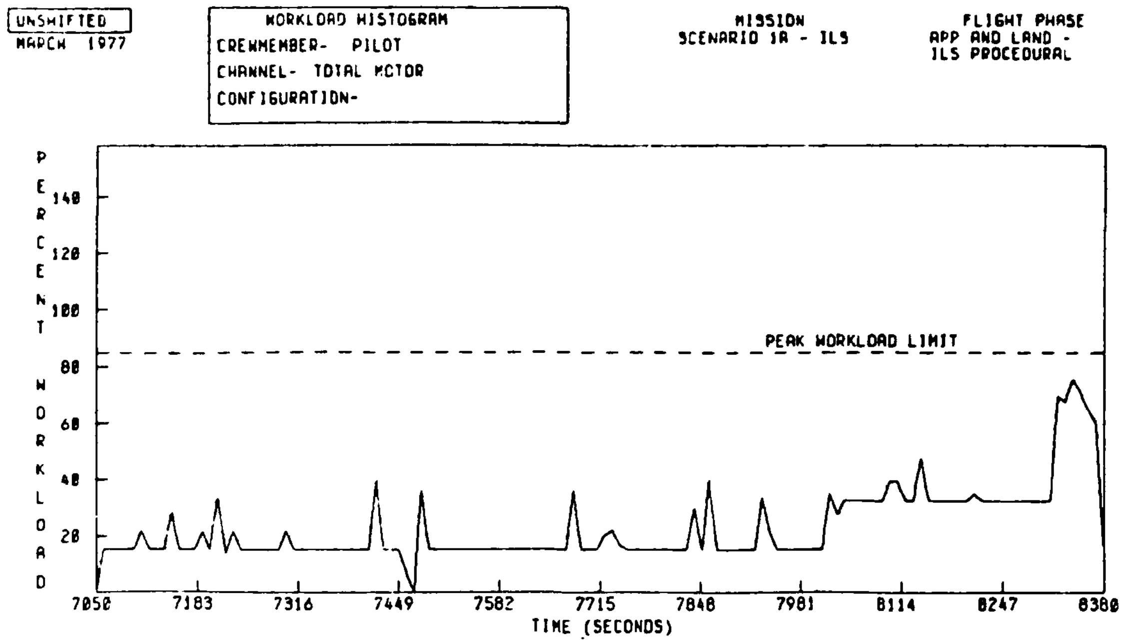


Figure 35 *Total Communications Time History*

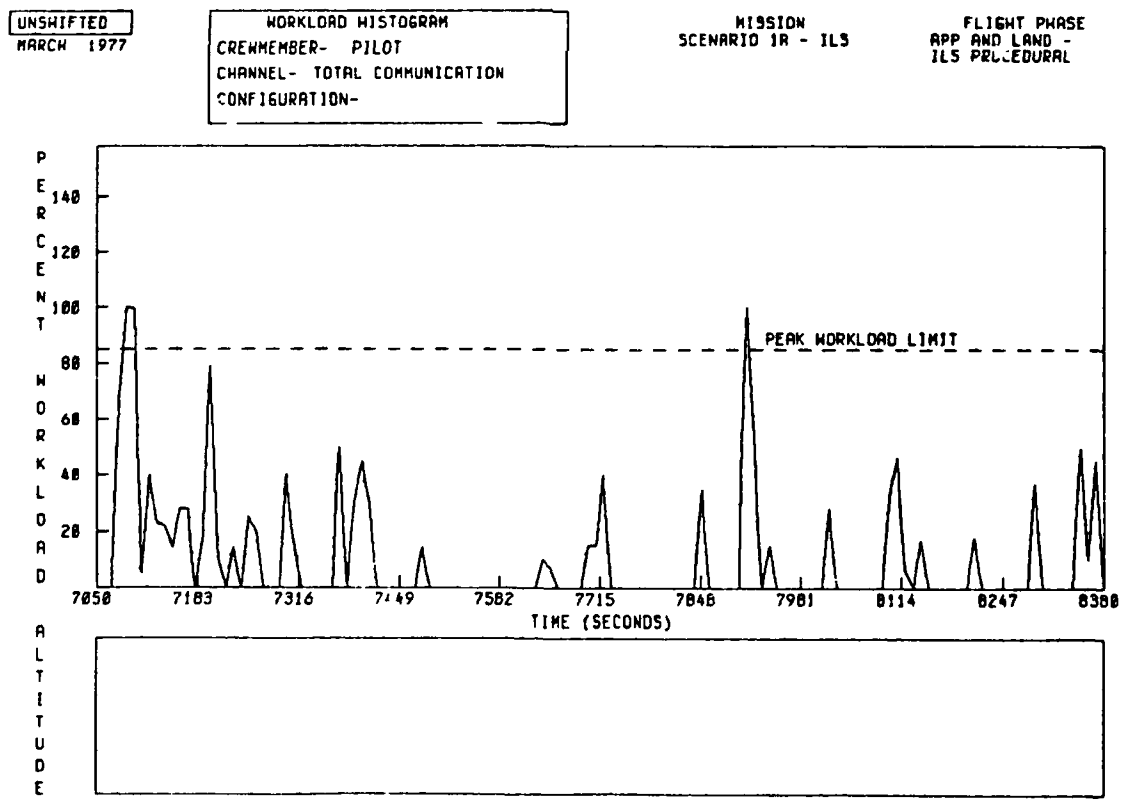


Figure 36 *Total Motor Time History*

The time required to cover the full visual field is a factor in traffic detection. If an aircraft not on a collision course enters the visual field of the crew, it will remain only momentarily in the visual field and then be gone. If the total search field can be reduced for each crew member, as would occur with an ATC advisory, then the smaller area (for example, 11:00 to 1:00 o'clock) to be scanned means that the target will probably remain long enough for the crew members nearest the forward windows to have sufficient time to scan that portion of the larger visual field with intensity. In this case, the probability of a contribution to detection by an observer with a more distant eye position is likely to be small, according to a study conducted in connection with workload evaluation on the B-737.

It should be kept in mind that the amount of time available for the pilot to look outside, particularly when there has not been an ATC traffic advisory, is not the same as the actual time spent looking. It is generally thought to be the case that people are not able to sustain longterm monitoring activities at an efficient level, using disciplined scanning patterns, for example, in the absence of reinforcing events. Aircraft detections at altitude tend to be made only when the stimulus is far above threshold, shows a long condensation trail, or when a special incentive to active search is present in the form of intercepted ATC communication or an advisory. Hence, unexpected detections may be rare events, and external search performance may often be perfunctory in nature. Such effects can hardly be determined by calculation or synthetic methods.

4.6 Other Computer Modeling Studies

In addition to the more general task-time analysis programs, which can provide summary workload histograms for flight scenarios (with percent workload plotted against a time coordinate), average times for each activity category, and other summary presentations, individual manufacturers have developed specialized computer programs for various aspects of crew workload evaluation. A wide inventory of programs have been reported in the literature, and aircraft manufacturers have exercized several of them to the point that they have knowledge of various uses. Examples include programs used to calculate geometric data used to evaluate workload due to eye angles and linear distances, and programs used with eye/hand motion and procedure execution times to provide quick-look workload estimates. As indicated in Sections 2.3 and 3.4, these tools are appropriate at various early stages of flight deck design. A possible contribution to workload documentation after completion of design may be demonstrated in the future.

4.7 Summary Comparison with Reference Aircraft

The final determination that the design flight crew can operate the aircraft safely has been a critical issue only in the case of aircraft designed for fewer than three crew members. An active debate on crew composition has continued in one form or another since World War II, and over that time period flight crews of large aircraft traveling long distances have been reduced in stages. First, the specialized radio operator was eliminated when communications facilities improved. The navigator was found to be unnecessary when better ground aids and cockpit equipment made it possible to follow a planned course with precision without periodic

celestial updating. Finally, the reliability of modern jet engines and related systems called into question the justification for including a flight engineer or nonflying pilot. With older piston aircraft and complex fuel transfer systems, the flight engineer had important duties. In the newer turbojet aircraft with increased automation, there is less for a third crewmember to do.

Under today's regulations, to approve a minimum crew size of two, the question that must be settled is the suitability/acceptability of the workloads on the Captain and Co-pilot and the demonstrated ability for one crewmember to fly the airplane suitably in a high denisty, high workload environment. If workload demands on either of them are excessive, or if it appears that a third crew member can assume some of the tasks that are causing specific problems, then three crew members may be required. If, in contrast, it is shown that all necessary operations can be conducted in a safe manner by the first two crewmembers, that crew size will be deemed adequate. Hence, the critical facts that must be submitted to the TCB are those establishing the ability of the design crew to conduct all anticipated flight scenarios with acceptable workload and attainment of safety.

The important role of comparisons of individual crew members workload and total required procedures between the new and reference aircraft is to show that the new design does not have increased task demands in a quantitative sense. The overall adequacy of flight deck accommodations, comfort, and reasonableness of task demands in a qualitative sense, are also assessed separately from any task/motion or other computer model type calculation. No method is known to make such overall qualitative assessments other than objective, unbiased pilot judgment.

These qualitative assessments by company test pilots have been proven to have considerable value since accountability for the individual pilot judgments has a significant influence on the reliability of the assessments offered. Also, as indicated several times, there is a strong tendency to be conservative in the number and extent of changes in the cockpit of a new aircraft design; therefore, the probability for error in the qualitative assessment is held to a low level.

For significant departures from the time/operationally proven systems of earlier aircraft types, more extensive qualitative evaluation of the task demands is required. Where these departures from the use of proven systems are contemplated -- such as the application of new technology -- and no suitable evaluation in simulation and flight of related military applications has been accomplished, a dedicated technology application development program is conducted by the manufacturer(s) -- sometimes with NASA or FAA sponsorship. The resulting programs usually involve development and evaluation programs spanning several years and include evaluations conducted in both simulation and flight. The evaluators associated with these test programs, include experienced test pilots employed by the manufacturer, customer airline pilots and, depending on the sponsorship of the program, may also include government test/evaluation pilots (i.e., NASA, DOD and FAA pilots). The development of the advanced electronic display system for the SST was an example of the latter type program.

5.0 Procedures - Flight Test

5.1 Functional and Reliability Test Flights

The final crew workload studies made for certification are normally part of the functional and reliability testing (F&R) of the airplane, although in some recent cases of 2-crew airplanes, there have been additional simulated airline tests appended following completion of F&R testing. These tests are made with direct FAA participation.

The manufacturer prepares a plan to provide realistic in-service conditions so that actual workload can be observed, recorded, and reported for both normal and abnormal (contingency) conditions. Critical elements of that plan, which must be approved by the FAA, include: the selection of test pilots, the number of test crews, the selection and preparation of flight crew observers, the data recording methods to be used in-flight, the data analysis procedures to be applied during and after flight, the particular contingency conditions and combinations of conditions to be covered in the flight tests, the number of replications of tests, the route structure, length of duty cycles and stresses to be controlled in the flight environment (include weather and ATC problems), and the criteria of success. In addition to the initiative required of the type certification applicant to prepare the test plan and the FAA participation that is aimed at ensuring that the plan is adequate in all respects, it may be the case that outside parties will make proposals for additional or special test conditions. All such proposals will be considered jointly by the manufacturer and the FAA, but such outside contributions of ideas and information create no obligation on the part of the FAA to discuss the overall plan or its details with such informants. The sole responsibility for creating and approving the plan rests with the manufacturer and the FAA.

As discussed in Sections 2.3, 3.1, and 3.5, flight test is a continuing process, not a phase initiated and completed as part of aircraft certification. The F&R program is such a phase of airworthiness proving, but represents the culmination of a larger testing activity, not the whole substance of that activity. Ordinarily, new developments -- particularly new system concepts -- are flight tested for as many as five years or more before going into a new aircraft program. Not only are the potential safety issues identified and resolved during this extensive lead time, but the economic viability -- cost, reliability and customer acceptance -- must be established before aircraft manufacturers will risk investment for new technology applications in critical aircraft systems for a production aircraft. Examples of these long-term development programs include: the over-ten year old FAA/Boeing advanced digital instrumentation and control system currently flying in the NASA-Langley TCV airplane; numerous flight test programs on advnced cockpit displays, flight management systems and avionics developments conducted by the major avionics manufacturers; government (DOD, FAA and NASA) flight tests of various new developments in navigation, landing, and control systems; and numerous in-house developments conducted by the airframe manufacturers. This substantial flight development and proving often leads to the development of industry-wide standards, in many cases with promulgation by AEEC, SAE, and similar

broad-based committees of published "Specifications and Characteristics," issued by ARINC, "Aeronautical Recommended Practices," by SAE, "Technical Standard Orders, " (TSO's) by federal agencies, and the like. These independent technical organizations have committees involved in the development of standardization criteria for virtually every aspect of transport category airplane system and cockpit feature (i.e., aircraft handling qualities, display considerations, visibility requirements, lighting, alert and warning systems, instrumentation grouping, environmental, operational terminology and system reliability). The membership is composed of a broad spectrum of the aviation community -- air carrier, airframe manufacturer, avionics manufacturer, pilot unions, government, etc. -- and includes international participation. Thus, before a new development is injected into an aircraft program, where there would be a risk to the certification applicant if the system proved unsatisfactory in actual flight operations, both extensive flight test and approval of design, manufacturing, and interface and standardization criteria normally have been achieved.

In view of the above, certification flight tests do not constitute the experimental, comparative, and procedures evolution phases of new flight deck system development. Rather, certification represents a brief but systematic examination and confirmation of the configuration and particulars of integrated function. The component equipment either has been proven in previous aircraft or, if new, has been proven in long-term flight test programs. Certification testing proves that the cockpit systems integration has been done right so that the crew and all systems can work together.

5.2 Schedules for High Workload Flight

In recent certification programs for two-pilot crew aircraft, eastern routes including Washington, New York, and Boston have been used with duty-day length and number of operating cycles patterned on the longest and busiest found in any airline schedule. To further increase the potential for fatigue, individual crews have been scheduled to fly the test aircraft on as many as three consecutive days. Concentrated programs of flight test in high density environments have been used to simulate severe airline operating conditions.

In the B-737 certification program, the heavy traffic, noise abatement procedures, the sawtooth altitude approaches to clear Kennedy traffic, the holding patterns, and short cruise times made these routes ideal for the test. Conducting the test over a full week during the Thanksgiving holiday period with flights scheduled over a twelve-hour period was selected to assure peak traffic for at least part of the test period. The weather provided both dry, clear days, and cloudy, foggy and wet days for the flights. Furthermore, crew duty times were increased beyond normal to provide opportunity for replication of the most unfavorable pilot schedule conditions.

Fifteen flights of the test were conducted with simulated inoperative equipment covering almost all malfunctions that could significantly contribute to high workload, such as autopilot failure, cabin pressurization

failures from auto and standby modes, generator failures, and hydraulic failures. Single pilot flights were conducted from each pilot's seat. The combination of the holiday period, weather, day and night, flying eastern routes, simulated malfunctions, single pilot operation, and a pilot crew with almost no previous experience on the routes provided a thorough, comprehensive and extraordinarily stringent test for two-man crew evaluation.

In the case of the stretched DC-9, model 50, five days of daytime flights and one day of night flights were conducted in the Yuma and Los Angeles areas, with IFR and VFR conditions, different types of landing approach (coupled, manual, and circling) and selected contingency conditions (pilot incapacitation, MEL items such as generator out, and emergencies such as engine failure on approach, pressurization failure, and unruly passenger). Three crews flew the 8-hour day schedules, representing the manufacturer and the FAA. With the later DC 9-80, an eastern route segment was selected between Atlanta and Boston.

Contrasting with the airline type scheduling for high workload flight testing, various recent three-crew aircraft have used schedules that were less structured but none-the-less stringent. The B-747 was taken on a world tour, for example, to examine crew performance under highly varied workload conditions produced by a range of ATC, airport aids, and traffic environment factors. While the B-747 tour preceded certification, the B-727 was taken on a world tour after certification. At the time the B-727 was put through F&R testing, its three-crew flight deck raised few critical workload issues, so the F&R flight test program was pointed to provide a sufficient service experience to establish the reliability of operations with a particular layout of aircraft systems, each of which had been extensively tested in other operations.

5.3 Selection of Test Crew Members

The selection of experimental test pilots has been one of the most controversial aspects of the workload flight plan. To date, all experimental test crews have been made up of FAA pilots and manufacturing company pilots. Typically, these crewmen are highly experienced experimental training pilots as well as qualified engineering pilots, so they are well aware of the factors likely to affect the least proficient line crew. However, it is recognized that such pilots differ from typical line crews, to whom it is intended that the workload findings should apply, in at least three possible ways: they are less familiar with routes, individual airports, and conventional procedures in the test area than would be line crews, but counterbalancing that unfamiliarity disadvantage, these test pilots may be more disciplined and orderly in planning and execution of flight duties; finally, the company and FAA pilots may be motivated differently than some line crews. Hence they may be less or more affected by the special stress of test conditions with continuous recording and special observers. Many times the question of including typical line crews in workload flight tests has been raised, but to date no alternative solution has been found to the problem of how to ensure objectivity and representativeness. Although not stated in FAR 25.1523, it seems obvious that workload test results are intended to cover the situation of the least proficient, most fatigued, and least coordinated

line crew. But no idea has been developed as to how to select such a worst-case, non-hazard-involving sample, or if selected, how to ensure realistic performance. F&R flight tests are expensive in a cost-benefit sense, and, when injected into a busy traffic environment, add a burden to the overall system. Hence, it has been unfeasible to select crews at random in hopes that a large sample would include all possible variations in skill, crew interactions, and unique reactions. However, experience so far has been good; an adequate solution appears to have been demonstrated in recent models.

By a cost-benefit sense, what is implied is that workload tests in F&R flying have not produced new information of great value. Starting with flight number one by the first production aircraft, participating pilots provide comments on workload and on other aspects of the flight operation. These observations build up prior to the formal test period in F&R, when there occurs a condensation of a mass of workload observations for the record. The sample of pilots, overall, flying the aircraft and having opportunity to comment is large, and in this sense, workload evaluation is a part of all flight tests. It would be remarkable, indeed, if a flight deck configuration was accepted by this larger sample, under widely different test conditions, and then attained an alternate rating of approval in a special F&R test unit.

Rather than use a broad sample, or a pinpoint selected smaller sample, the solution to the problem of deciding who should fly the new aircraft in workload tests has been found in conducting the F&R flights with major participation by pilots who are experienced in workload testing and are knowledgeable about high workload situations in other well-proven aircraft types. Regional flight test people have had continuity from one certification program to another and qualify as "trained observers" for purposes of determining that the new aircraft is easier/harder/equally hard to fly, conduct required procedures in, and solve problems in.

An occasional recent practice has been to employ outside consultants to confirm workload results. Such consultants would normally be selected by the FAA from special pilot populations, such as those with exceptional breadth of flight experience or special engineering/technical background. These consultants normally would have extremely limited time in the type aircraft under test.

It is important that all test crews have adequate ground school and simulator training in the features, check lists, and flight manual of the test type prior to conduct of data flights. Normally, several of the pilots used in F&R workload tests will have additional experience with the new design gained during the aircraft design process and earlier, non-data test flights. Those pilots may also be familiar with the outcome of simulator and computer evaluations of workload using the new flight deck design and procedures. This prior familiarity gives those pilots an advantage in making the inflight evaluations since they know what especially to look for, what are the *pro forma* advantages of the new design, and which procedures have been changed due to new equipment, simplifications, or added automation.

In the case of the two-crew aircraft, the usual number of test crews is not less than three teams, each consisting of company and FAA pilots. Each crew is expected to operate the aircraft for a period of time sufficient to reproduce line-service fatigue conditions and also to serve as observers and evaluators of other crews flying on subsequent days. Hence, the "duty" periods of the test crews consist not only of the days that they operate the test aircraft, but added days riding jump seat or observing flight deck activities on video monitors from the cabin.

5.4 Objective Data Recording Inflight

If it were possible to do so, pilot workload would be measured objectively, i.e., by collecting and analyzing data representing actual measurable events, pilot actions, and flight outcomes. Such objective measurement would avoid the pitfalls of introspective reporting, which may include individual bias, incomplete self-knowledge of physical events, and errors of recall.

Many of the test procedures described in previous sections of this report do produce objective data through direct measurement of the frequency and duration of observable actions and the adequacy of flight control performances. But none of these objective or semi-objective methods produces an overall estimate of total workload comparable to the subjective assessment that can be provided by either the pilot flying the aircraft or a qualified observer. Instead, it is part tasks and external action aspects of total workload that are measured objectively, while the cognitive workload of planning, thinking, and problem solving is only partly covered. It may be at some future date that measurements can be taken that indicate what is happening in the mind of the pilot, but proven procedures for doing this in a total workload sense are not presently available.

At the present stage of evolution of proven methods of evaluating crew workload, procedures that have some "objective" and some "subjective" aspects are still necessary. While we know that it is unsatisfactory to rely entirely on the pilot's rating of his own performance, we also know that a highly trained pilot observer who knows the pitfalls of self-rating can note much that is normally missed by the self-rater. Also, filling a function like the umpire in a sport, the trained observer can fault a performance even though the actual crew would give it a success rating. This is the role provided by training pilots. The priority, then, must be put on the assessments made by "senior" observers playing the part of a referee. To support their analysis of the flight events, video playback of flight deck signals and pilot actions has proved of great value. An overall "good" rating covering an 8-hour flight with various departure cycles would be expected to be less valuable than a phase-by-phase or contingency-by-contingency evaluation made after observing the entire flight and then replaying the records of various high workload phases. Discussions among several such trained observers while replaying the pictures and questioning each other as to the reasons that the crew acted as they did can be valuable in adding insight into the ratings.

In the B-737 workload program, video was used primarily to present a picture of flight deck events to observers situated in the rear cabin making real-time evaluations. In the later DC-9 programs, recordings were made with the principal intent being later review and analysis. When there is a particular cockpit feature that is changed or a new system that is added, video tape may be employed specifically to record the use and performance of that feature. Somewhat related procedures were used in the B-747 program. Tests pilots asked for review of specific flight deck procedures after seeing a film that was made in a mockup of all usual procedures. Using the mockup, these pilots then reviewed the questioned procedures and repeated the necessary actions in live action to recheck accessibility, visibility, and related concerns. Also on the B-747, what was called a "Substantiation of Analysis by Live Demonstration" was conducted by putting the pilot through a structured simulator flight with tape pacing of events. This constituted one form of a "live" time-line.

Objective data to be recorded during test flights normally includes aircraft performance parameters such as altitude, speed, bank angles, and the like, the ATC record of clearance changes and other communications, and the actually attained schedule of planned contingency events plus such unplanned outages and problems as may have occurred. Using the method of Douglas Aircraft Company, analysis of tapes made during flight takes 1-1/2 hours coding time for each minute of actual flight. The result of that labor-intensive study of small segments of crew activity is a detailed record of every observable crew member action arrayed in a chronological sequence showing when the event began, how long it lasted, and how often each action was repeated. An objective record of this kind, representing, for example, a segment of 10 minutes of an arrival in a high density environment, is used in two ways. First, it gives the trained observer a solid basis to reconstruct the pilot reactions and altered procedures that were forced by the contingency conditions employed. This tends to ensure that the observer's rating are as "objective" as possible. Second, the record based on events during actual flight can be compared to the similar workload estimates generated in simulation. This comparison may permit an estimate of the validity of the measures that were based on computer models or part-task simulations.

Two sample detailed records of objective data recorded inflight during flight testing of the DC-9-50 are illustrated in Figures 37 and 38. The first shows altitude, air speed, and heading relationships during the 15 minutes prior to landing. The second shows a time record of several activities of the Captain and First Officer, including communications, visual activities, and hand actions.

Other techniques do support such analyses and are being developed. For example, in recent years, development of improved oculometers and pupillometers has made it possible to begin to record pilot eye fixations. Because of the wide range of eye and head movements, it is estimated that eight oculometers would be needed to fully instrument a cockpit. Obviously, this cannot be done in a flyable aircraft. With one recording device, a particular area of the panel such as the flight instrument basic T can be handled, and some data of this kind is becoming available from ground simulator tests. From the NASA-Piedmont Airline study, some question has

been posed about the scanning patterns employed with the T arrangement, and this raises an interesting question. From experience it is known that aircraft having flight instruments arranged in the T have good records. If simulator data should suggest a change in that instrument arrangement, without support from pilot assessments made in actual flight, what would be the appropriate action?

5.5 Evaluation By Pilots and Observers

From time to time, criticism has been voiced about substitution of the one readily available method of assessing total pilot workload and comparing it to that of another aircraft, pilot "subjective" evaluation, for the more quantitative and methodologically desirable "objective" measurements. Workload authorities make two points related to this criticism. First, the history of past workload determinations by pilot assessment shows that the correlation between the informed judgments of highly qualified and trained experimental pilots, on the one hand, and the proof obtained in extended periods of diversified line experience, on the other, is essentially perfect. Second, there is no other comparable human team performance that can be rated in toto by so-called "objective" methods. It is a fact of life that there are strict limits to what we can measure by directly instrumented procedures, and these limits never include all important cognitive activities and emotional experiences, such as creative problem solutions, subjectively experienced stress, and both reactive and "free-floating" anxiety. Hence, it may be foolish to overstress the need for objective measures per se. Rather, the goal that is more appropriate may be to "objectify" pilot assessments by providing appropriate structuring to both the experience and recall aspects. By structure, it is implied that systematic exposure will be given to the relevant conditions, so that the pilot-rater has immediate experience with the work situations that are to be assessed. Also, the recall situation can be structured by use of standard inquiry procedures and aids to recall such as playback of recordings and provision of the part-task objective records.

In the B-737 flight tests a combination of 14 FAA observers from both Washington, D.C., and the Western Region, air carrier, air traffic control, and engineering, as well as observers from the Boeing Company, engineers from the 737 Control Cabin Equipment Group, and 737 Flight Test Pilots, made up the evaluating technical team. Television and audio of all cockpit activities were provided in the passenger cabin area to provide all the observers with full information of cockpit action. The video, audio, and basic flight parameters were all recorded to provide means for full playback as well as on-the-spot evaluation.

Every means had been taken to provide realistic in-service conditions so that actual workload could be observed for both normal and abnormal conditions. In addition, the video and audio portions of the flight were brought back and analyzed. As each flight was viewed, a team of people recorded with stop watches the number of minutes the crew could be observed or heard to be taking action in each of 6 categories:

Navigation

Consulting maps and charts and setting RMI, course, flight director, and autopilot.

Communication

Sending or receiving communications

Radio Tuning

Tuning radios for either navigation or communication purposes.

Controls

Actuation of flaps, gear throttles, etc.

Systems

Operation of fuel, hydraulic, electrical pressurization, anti-ice, etc. during flight.

Miscellaneous

Includes all other observed cockpit actions needed in flying an aircraft such as check list reading, writing down clearances, etc.

The resulting figures showed the time spent in climb, cruise, descent, and approach that the crew members used for each of the above categories.

The remaining time was available for outside visibility and handling malfunctions. A word of caution on these test results; all tests have their limitations and this is no exception. Two of the major things that were not shown in this crew action study results are: (1) control movements that were so small as not to be detectable, e.g., small corrections in control wheel steering on the B-737; and (2) the multiple tasks such as communicating while still maintaining outside observations. Of course, the previous disclaimer covering covert activities also applies; thinking, problem solving, planning, and feeling could not be assessed from the records.

At the end of a day of F&R flight that includes workload evaluation, it is usual to ask each crewmember and designated observer to complete a standard form covering special features of the day's flight, such as weather and turbulence encountered on each leg of the trip and the type of approach conducted at each terminal. In addition, there is usually a more detailed record form that tallies specific problems such as completeness and correctness of all check lists, notation of necessary check list interruptions, determination of the acceptability of crew workload upon special events such as a last minute change in runway assignment, special air traffic routing, specific equipment outages, crewmember

incapacitation, and the like. These record forms are intended to assist the raters in covering all critical aspects of the flight and, thereby, avoid any "halo" effect as might occur when an overall successful operation causes one to forget some detailed aspect of the flight that was not satisfactory.

When these tabulations are completed on the day of flight, an analysis of the objective data will not yet be available. Prior to final acceptance by the TCB of the summed evaluations by pilots and observers, at least some sampling of the key objective records will have been made to ensure that tabulations made from event records provide supporting evidence for the pilot and observer evaluations.

Each participating pilot and observer is asked also to state his rating of overall flight workload due to air traffic control procedures, communications, navigation, and collision avoidance and to rate the workload level for all procedures combined. Very recently a three-attribute rating of pilot mental workload has been developed to assist the assessor to explain observed workload in terms of the fraction of time busy, the intensity of required mental effort, and the intensity of emotional stress (see reference 22). This procedure uses a descriptive system patterned on the Cooper-Harper aircraft handling quality scale, with three choices under an overall rating of satisfactory, three more under an overall rating of acceptable, three more under an overall rating of unacceptable, and one for impossible. This new three-attribute scale has not been used in certification programs to date, and is reported here only as an extension of previously employed rating techniques. For past successful aircraft certification programs, there has not been a single, standard method of collecting overall pilot and observer workload evaluations. In every case, however, the participants have given both overall and specific ratings of the acceptability of workload encountered during the actual test flights. The key guideline here is that all the data necessary to substantiate the design must be collected. If the panel of expert pilots concurs that the flight deck design is acceptable and compares favorably with that of proven inservice aircraft, this substantiation has been obtained.

One perspective on the flight test demonstration of workload acceptability includes the idea that from the start, the FAA is looking at progress with certain milestones. At the point of conduct of F&R testing, or similar inflight replication of major features of ordinary airline service, the question is asked: Is this airplane tested and described well enough so that it is reasonable to write flight scenarios that are adapted to its special features and that will give the opportunity for any novel properties to be assessed? At completion of the review of all certification data and the decision point as to actual award of the airworthiness certificate, the question has evolved to: Is this aircraft ready to be put in service with regular line crews?

Before actual revenue flying is initiated, the line crews are trained and route checked. This process is closely monitored because several important steps besides crew qualification are taking place. First, the airline

is writing its own procedures for standard crew functions in the new aircraft. Secondly, the airline is installing items of its own equipment. Finally, an MEL is approved for the particular operation. Any of these may impact the overall configurational suitability of the flight deck design; hence, final determination on this point, and control of the final line-service use of the aircraft is vested in the FAA regional office having regulatory authority on that particular airline's operations.

This means that, for all practical purposes, the FAA chief inspector for each airline putting the new aircraft in service has an additional review of the operational acceptability of the new flight deck, and he makes this review with specific knowledge of the procedures, equipment, and operating limitations that may be applicable, individually. Since this is a review by highly experienced operations inspection personnel, pilots familiar with the airline's pilot training and qualification programs and any operating problems that may have been revealed in earlier line service over the same routes, an important degree of assurance is obtained.

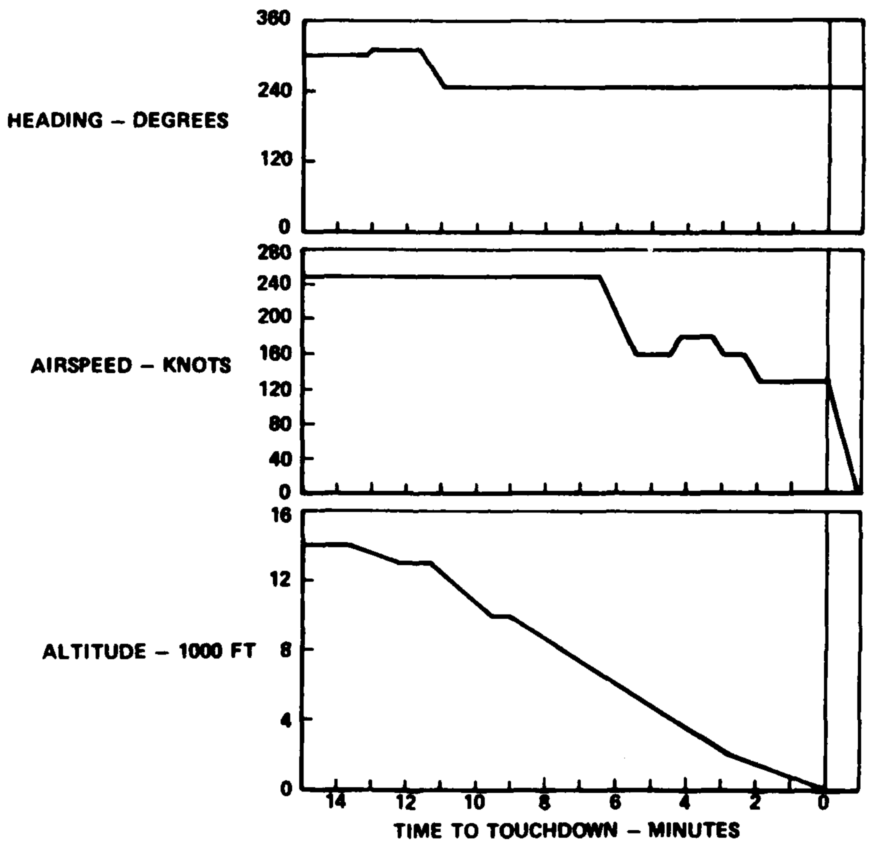


Figure 37 Altitude, Airspeed and Heading Relationships

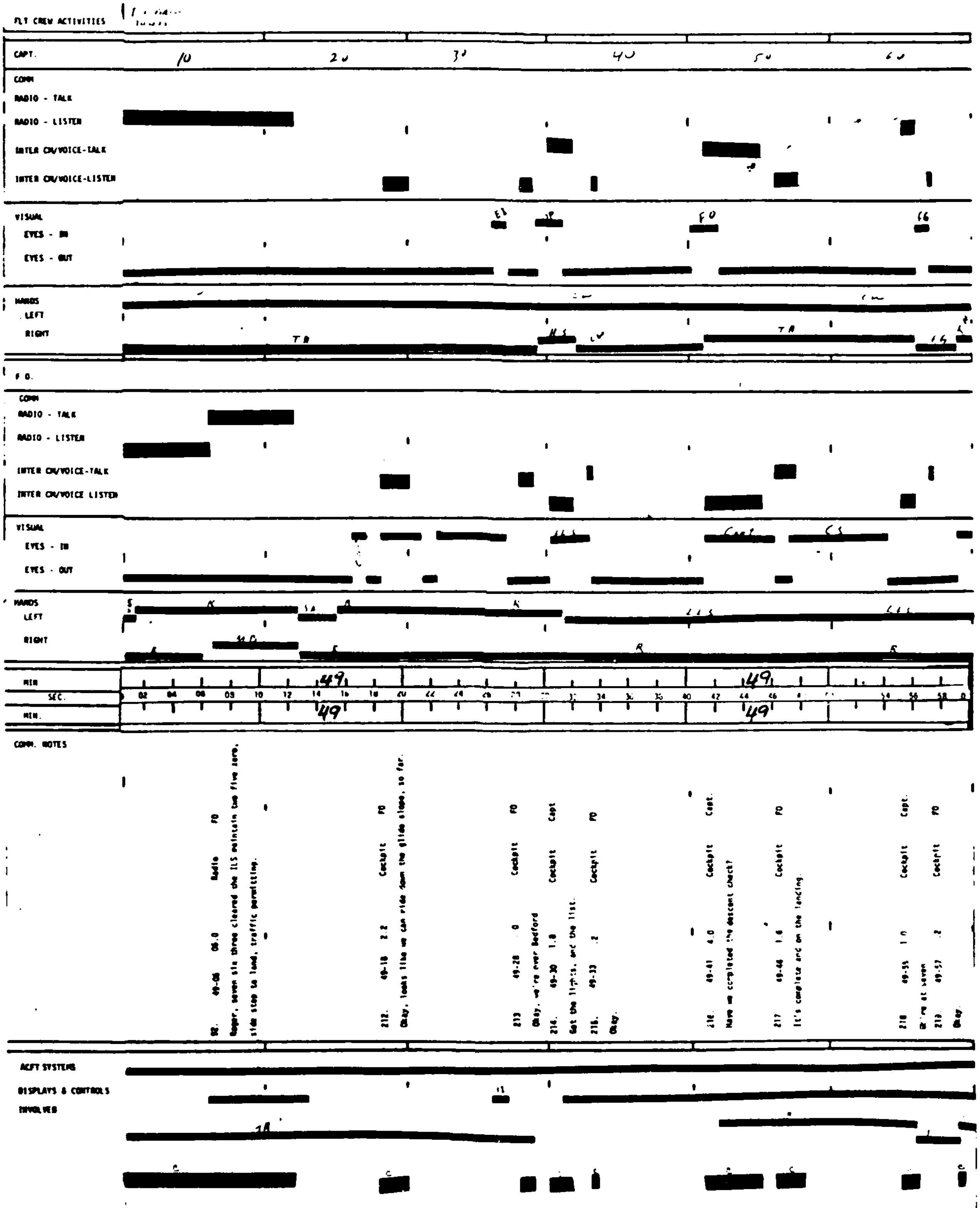


Figure 38 Flight Crew Activity Vs. Time Record

AD-A114 167 WRIGHT STATE UNIV DAYTON OH SCHOOL OF MEDICINE F/G 1/3
FLIGHT CREWMEMBER WORKLOAD EVALUATION.(U)
APR 81 R L SULZER, W J COX, S R MOHLER F33615-80-K-3627
UNCLASSIFIED DOT/FAA/RD-82/21 NL

2 OF 2
AD
A114167

END
DATE
FILMED
5-82
DTIC

APPENDIX A

This appendix contains examples of letters and correspondence between an applicant for certification of a single aircraft and the FAA.

AUG 22 1980

B-7673-EA-14900

Department of Transportation
Federal Aviation Administration
FAA Building, King County Int'l Airport
Seattle, Washington 98108

Attention: Mr. Charles C. Schroeder, Chief - ANW-210
Engineering and Manufacturing Branch

Subject: Letter of Program Definition for
Qantas (QAN) 747-238B, RD533,
Follow-on Combi

Reference: (a) Project T2245NW-D
(b) Boeing letter B-7673-EA-13056
dated May 3, 1979

Gentlemen:

This letter introduces a program for the certification of a follow-on Qantas 747-238B Combi airplane.

General Information

The airplane is scheduled for delivery in October, 1980 and is identified as follows:

Model	S/N	Line No.	Block No.	Registration
Combi	22615	483	RD533	VH-ECC

The airplane will be powered by RB211-524B2 engines and certificated at the following gross weights:

Maximum Taxi	823,000 lbs.
Maximum Brake Release	820,000 lbs.
Landing	630,000 lbs.

The airplane is identical to RD532 certificated under the reference project in November, 1979 except for minor production improvement changes which do not affect the performance, handling, or operational characteristics and the following differences.

Structures

The RD533 airplane will be a twelve-pallet combi while RD532 was a six-pallet combi. There are, therefore, some minor structural differences in the provisions for 833,000 lb. maximum brake release gross weight.

Noise Certification

Noise data substantiation will be the same as that provided for Qantas RD518, Project T2567NW-D.

Automatic Flight Controls

These will be the same as on Qantas RD518.

Supplemental Air Circulation Fans

Three additional recirculation fans will be installed in the lower lobe to provide ventilation when operating the air conditioning packs in a "fuel saving" mode. This mode is provided by pack flow control valves which have a "half-flow" setting. Filters are installed upstream of the three recirculation fans. An electrical inter-tie will automatically shut down the fans when the cargo fire extinguishing system is armed. The galley/lav vent system will be modified so its air is exhausted through the outflow valves in flight.

The controls are on M170 on the flight engineer's panel and are depicted on the enclosure to this letter.

The reference (b) letter transmitted D6-13333, Appendix M, Description and Failure Analysis for Economy Mode Air Conditioning with recommended approval by A. B. Hartley.

It is proposed that this installation be evaluated by DER during production flight testing.

Interior Arrangement

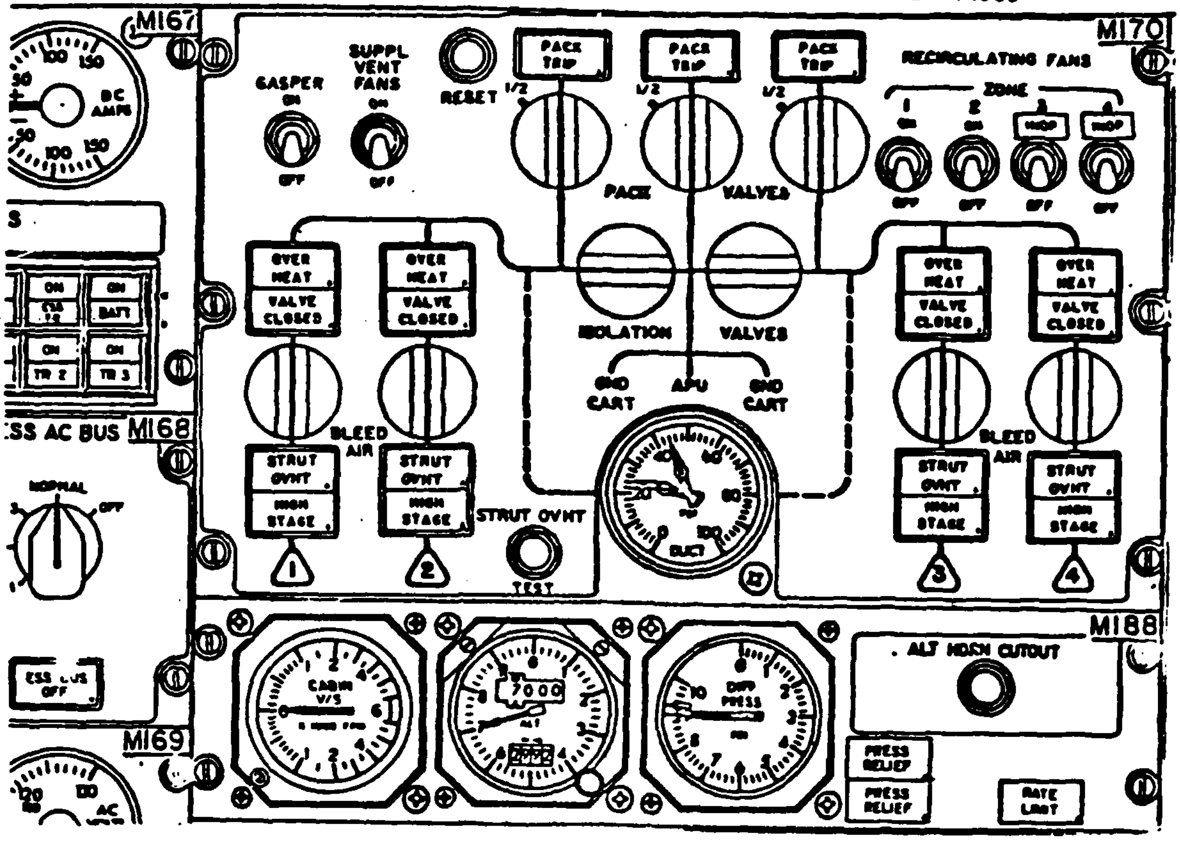
In Zone B of the main deck, the business class seating pitch will be increased. A minor interior inspection will be required.

Airplane Flight Manual

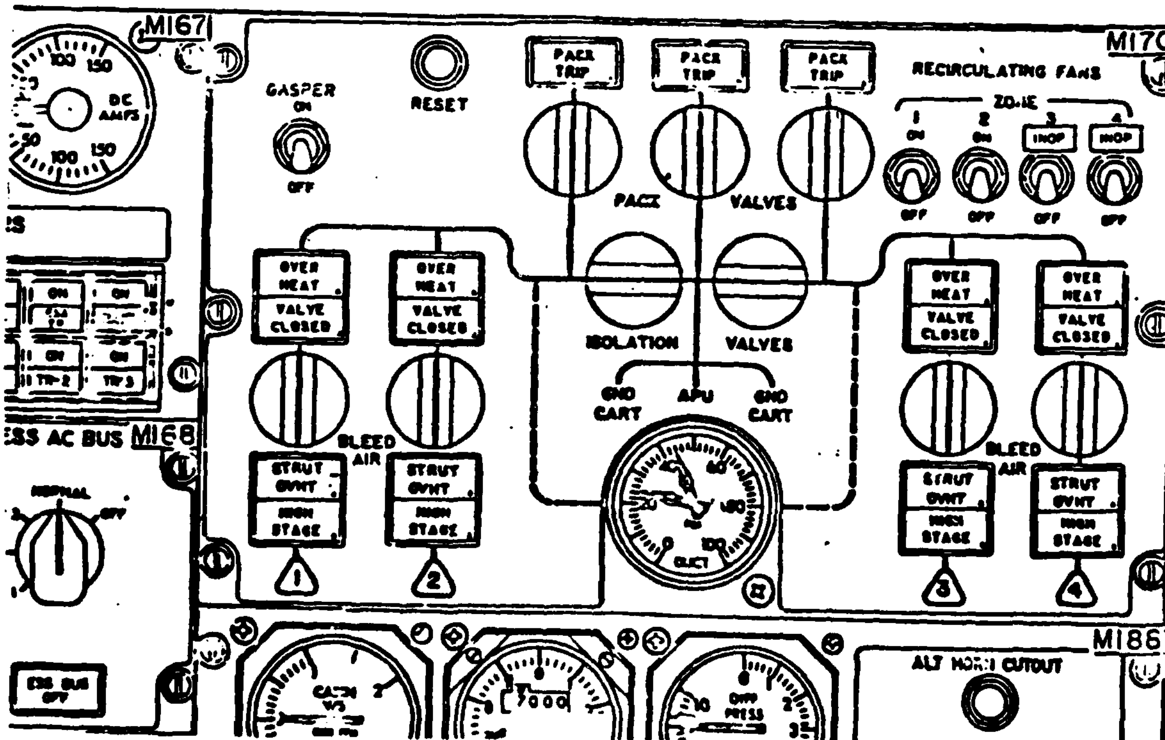
An Airplane Flight Manual revision will be submitted for approval.

Enclosure to
B-7673-EA-14900

FLIGHT ENGINEER'S P4 PANEL



RD532



Significant Differences Drawing List (SDL)

Systems and Structures SDLs are enclosed.

Testing

A flight test specification for DER evaluation of the supplemental air circulation fan installation will be submitted.

FAA Approval

Approval by DER according to Boeing/FAA Memorandum of Agreement dated October 23, 1978 is recommended.

A draft NW Form 8110-1 is enclosed.

Very truly yours,

THE BOEING COMPANY

ORIGINAL SIGNED BY

C. D. Engebretson
for Reginald Utting
Manager, Airworthiness
747 Division
Boeing Commercial Airplane Company

Enclosures

cc: L/D Distribution
P. J. Conlon

JRT:gp

BOEING

Enclosure to
B-7673-EA-14900

DEPARTMENT OF TRANSPORTATION
Federal Aviation Administration
Northwest Region

AIRCRAFT CERTIFICATION ELIGIBILITY

All findings of compliance with the Federal Aviation Regulations applicable to the 747-238B have been completed and notification is hereby given of related actions by this office:

1. Airplane Flight Manual ______________ applicable to this Model has been approved.

2. Type Certificate No. __________ is being revised effective this date to include the Model ______________.

3. Effective this date, production of the Model __________ will be authorized under the terms of Production Certificate No. __________.

4. Type Certificate Data Sheet No. A20WE, Revision No. __________ is being revised to include the following:

   ( ) The Model __________ under Section __________.

   ( ) A new Section for the Model ______________.

   (X) Serial Nos. eligible: 22615

   ( ) Applicability of the following exemptions: ______________.

   ( ) The following limitations specifically applicable to the Model ______________:

Date: ______________

Chief, Engineering and Manufacturing Branch
FAA Northwest Region

cc: AWS-100, NW-EMDO-41, Applicant: ______________,
ACDO Involved: ______________

Original to ANW-210D

NW Form 8110-1 (10/79) Supersedes previous edition

Enclosure to
B-7673-EA-14900

SIGNIFICANT SYSTEMS DIFFERENCES LIST
(QF Model 238B, RD 533)

Drawing No.	Title	Used On
69B46137-18	M/A Galley/Lav Fan ** (Module Assy)	65B46006-5022
65B46118-174	M/A Air Cond ** (Module Assy)	65B46006-5022
60B00026-81	Flow Control Valve-** Dual Schedule	65B40004-1 65B40004-2 65B40012-1
60B40148-4	Particulate Filter **	65B48250-1 65B48250-2 65B48250-3
65B48250-4	Fan Assy **	65B48250-1 65B48250-2 65B48250-3

** Ref. PRR 79162-1

SIGNIFICANT STRUCTURES DIFFERENCES LIST
(QF Model 23?B, RD 533)

Drawing No.	Title	Used On
65B15802-5	Stiffener *	65B15087-201
65B15802-6	OPP 65B15802-5 *	65B15087-202
65B15880-1	Frame Instl *	65B15087-201
65B15880-2	OPP 65B15880-1 *	65B15087-202
69B18885-1	Stiffener *	65B15087-201
69B18885-2	OPP 69B18885-1 *	65B15087-202
65B15092-24	Tie-Shear *	65B15092-1 65B15092-2
65B15092-117	Web-Mid, Upper *	65B15092-1
65B15092-118	OPP-117 *	65B15092-2
65B15092-119	Tie-Shear *	65B15092-1
65B15092-120	OPP-119 *	65B15092-2
65B15092-121	Tie-Shear *	65B15092-1 65B15092-2
65B15092-122	Tie-Shear *	65B15092-1 65B15092-2
65B15092-123	Chord-Outbd *	65B15092-1
65B15092-124	OPP-123 *	65B15092-2
65B15092-125	Stiffener *	65B15092-1
65B15092-126	OPP-125 *	65B15092-2
65B15092-127	Stiffener *	65B15092-1
65B15092-128	OPP-127 *	65B15092-2
65B15092-129	Tie-Shear *	65B15092-1

SIGNIFICANT STRUCTURES DIFFERENCES LIST (Cont.)
(QF Model 238B, RD 533)

Drawing No.	Title	Used On
65B15092-130	OPP-129 *	65B15092-2
65B08315-49	Tie-Shear *	65B15092-1
65B08315-50	OPP-49 *	65B15092-2
69B18892-1	Shear Tie *	65B15088-161 65B15088-162 65B15089-1 65B15089-2
69B18892-2	Shear Tie *	65B15090-1 65B15090-2 65B15091-1 65B15091-2
69B18892-3	Shear Tie *	65B15090-1 65B15090-2 65B15091-1 65B15091-2
69B18892-4	Shear Tie *	65B15090-1 65B15090-2 65B15091-1 65B15091-2
65B08241-13	Strap (Bear) *	65B15087-201
65B08241-14	OPP 65B08241-13 *	65B15087-202
65B01893-235	Web *	65B01893-147
65B01893-236	OPP-235 * (Except As Noted)	65B01893-148
65B01893-237	Doubler *	65B01893-147 65B01893-148
65B01893-238	Doubler *	65B01893-147 65B01893-148
65B01893-239	Doubler *	65B01893-148
65B01893-240	Doubler *	65B01893-147 65B01893-148
65B10364-15	Chord *	65B01893-147
65B10364-16	OPP 65B10364-15 *	65B01893-148

SIGNIFICANT STRUCTURES DIFFERENCES LIST (Cont.)
(QF Model 238B, RD 533)

Drawing No.	Title	Used On
65B10097-31	Chord, Outbd *	65B15096-1
65B10097-32	OPP 65B10097-31 *	65B15096-2
65B15410-19	Skin-Machined *	65B15087-201
65B15410-20	OPP 65B15410-19 *	65B15082-202
65B15410-21	Skin-Machined *	65B15087-201
65B15410-22	OPP 65B15410-21 *	65B15087-202
65B15410-23	Skin-Machined *	65B15087-201
65B15410-24	OPP 65B15410-23 *	65B15087-202
65B15410-25	Skin-Machined *	65B15087-201
65B15410-26	OPP 65B15410-25 *	65B15087-202

* **Ref PRR 79272-4**

REFERENCES

1. Public Law 85-726, 85th Congress, S. 3880, August 23, 1958, 72 Stat. 731.

2. ANON: Federal Aviation Regulations, Part 25, "Airworthiness Standards: Transport Category Airplanes," Paragraph 25, 1523, Appendix D., Department of Transportation, Federal Aviation Administration, June 1974.

3. Hay, George C., House, Charles D., and Sulzer, Richard L., "Summary Report of 1977-1978 Task Force on Crew Workload," Report No. FAA-EM-78-15, Department of Transportation, Federal Aviation Administration, Washington, D.C. 20591.

4. Cox, William J., Gold, Fay, and Sulzer, Richard L., " Update of Summary Report of 1977-1978 Task Force on Crew Workload," Report No. FAA-RD-81-34, Department of Transportation, Federal Aviation Administration, Washington, D.C.

5. O'Donnell, J.J., "Procrastination + Lethargy = Tragedy," Air Line Pilot, Volume 48: No. 12, December 1979 pp. 6-8.

6. Sells, S.B., and Berry, Charles A., Human Factors in Jet and Space Travel, Ronald Press, New York, 1961, pp. 83-93, 129-131, and 234-235.

7. Glines, C.V., and Morris, M., (Eds), "No Compromise With Safety: The Crew Complement Question," 1980, Booklet published by Europilote and U.S. ALPA, 1625 Massachusetts Ave. N.W., Washington, D.C. 20036.

8. Article, "Boeing 757, 767 Avionics Design Frozen," Aviation Week & Space Technology, November 3, 1980 Volume 113, No. 18, pp. 203-207.

9. Report, "Improving Aircraft Safety," Committee on FAA Airworthiness Certification Procedures, National Academy of Sciences, 1980, Washington, D.C.

10. McFarland, Ross A., Human Factors in Air Transportation, McGraw-Hill, New York, 1953.

11. Morello, Samuel A., and Knox, Charles E., and Steinwetz, George E., "Flight Test Evaluation of Two Electronic Display Formats for Approach to Landing Under Instrument Conditions," Langley Reaserch Center, NASA TP 1085, December 1977.

12. Levison, William H., Baron, Sheldon, "Analytical and Experimental Evaluation of Display and Control Concepts for a Terminal Configured Vehicle," Langley Research Center, NASA CR-3270, July 1976.

13. Salmirs, S., and Tobie, H.N., "Electronic Control in Advanced Terminal Area Operations," AIAA 2-74, February 1974.

14. Munger, S.J., Smith, R.W., and Payne, D., "An Index of Electronic Equipment Operability -- Data Store," Report No. AIR-C43-1/62 American Institute for Research, Pittsburgh, Pennsylvania, December 1962.

15. Brown, E.L., Stone, G., and Pearce, W.E., "Improving Cockpits Through Crew Workload Measurement," Second Advanced Aircrew Display Symposium, Naval Air Test Center, 23-24 April 1975, pp. 370-376.

16. Armstrong, Larry, "DAIS Takes Center Stage," Electronics, February 6, 1975, pp. 76-77.

17. Bateman, L.F., "Flight Decks For Future Civil Transport Aircraft," Journal of Navigation, May 1977 pp. 207-219.

18. Baty, Daniel L., "Rationale and Description of a Coordinated Cockpit Display for Aircraft Flight Management," NASA TM X-3457, Ames Research Center, Moffitt Field, California, Nov. 1976.

19. McIver, Duncan, and Hatfield, Jack J., "Coming Cockpit Avionics," Astronautics & Aeronautics, March 1876, pp. 54-63.

20. ANON: Federal Aviation Regulations, Part 121, "Air Carriers and Commercial Operators of Large Aircraft," Paragraphs 121.627(c) and 121.303(d), Department of Transportation, Federal Aviation Administration, 8/30/80 and 12/1/78.

21. Stone,G., Regis, E.R., and Gulick, R.K., "Executive Summary: DC-9 Super 80/DC-9-50 Comparative Flight Crew Workload Study," Report No. MDC J8749, 1980, Douglas Aircraft Company, Long Beach, California, p 10.

22. Sheridan, T.B., and Simpson, R.W., "Toward the Definition and Measurement of the Mental Workload of Transport Pilots," Final Report - January 1979, Contract DOT-OS- 70055, Flight Transportation Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

23. Parks, D.L., "Current Workload Methods and Emerging Challenges," in Moray, N. (ed), Mental Workload, Its Theory and Measurement, Plenum Press, N.Y., 1979.

24. Parks, D.L., and Springer, W.E., "Human Factors Engineering Process Definition and Criterion Development for CAFES," Boeing Document D180-18750-1, Especially pages 120-154, NADC Contract N62269-74-C-0693, June 1975.

25. Miller, K.M., "Time Line Analysis Program (TLA-1)," Boeing Document D6-42377-S, NASA, LRC Contract NASA-CR-144942, April, 1976.

26. Chu, Yee-Veen, and Rouse, W.B., "Adaptive Allowance of Decision Making Responsibility Between Human and Computer in Multi Task Situations," IEEE Transactions on Systems, Man and Cybernetics, Vol SMC-9, No. 12, December 1979.